2011

Functional consequences of cytosine methylation in mitochondrial DNA catalyzed by DNA methyltransferase 1

Lisa Shock
Virginia Commonwealth University

FUNCTIONAL CONSEQUENCES OF CYTOSINE METHYLATION IN MITOCHONDRIAL DNA CATALYZED BY DNA METHYLTRANSFERASE 1

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

LISA SALE SHOCK
Bachelor of Arts, University of Virginia, 2003

Advisor: Shirley M. Taylor, PhD
Associate Professor
Department of Microbiology and Immunology

Virginia Commonwealth University
Richmond, VA
August, 2011

Acknowledgement

I would like to recognize a number of individuals who have played instrumental roles in both my scientific and professional life as well as my personal life throughout the duration of my graduate education. First, I would like to express my profound gratitude to Dr. Shirley Taylor for her continued guidance, patience, and friendship. Her unrelenting faith in my abilities, even as a "rookie" graduate student gave me the confidence to grow both scientifically and intellectually. Her approach to my training challenged me to develop an independence that has driven my productivity as a graduate student and will continue to benefit me throughout my career. I am extremely fortunate to have had the opportunity to work with her.

I would also like to thank the members of my graduate committee: Drs. Richard Moran, Tomasz Kordula, Todd Kitten and Francine Marciano-Cabral, for their dedication and for the invaluable contributions each of them has made towards my training. Throughout our interactions, each individual brought a unique perspective, challenging me to think critically and creatively and cultivating a scientifically well-rounded experience.

I would like to express my appreciation to Dr. Guy Cabral and the Microbiology and Immunology department for granting me the opportunity to train in the graduate program at VCU. I would also like to recognize Dr. Gordon Ginder and the Massey Cancer Center for their accommodation and support.

Thank you to all of the members of the Taylor laboratory, both past and present: Dr. Dolores Arjona, Erica Peterson, Prashant Thakkar and Joyce Balinang. I would also like to thank several members of the Moran laboratory, both past and present: Cortney Heyer, Dr. Scott Lawrence, Dr. Alex Racanelli, Dr. Scott Rothbart, Dr. Shane Kasten, Lin Xie, Guoyan Gao, Dr. Chen Yang, Stuti Agarwal, Catherine Bell, Audrey Thacker, and Will Buchwald. On a personal level, the camaraderie and comical relief provided by these individuals made the laboratory an enjoyable place to work, and fostered friendships that extended beyond the work environment. On a professional level, this group provided a degree of scientific and technical support that was vital to my training and success. I am deeply indebted to each of them.

I would like to thank my good friends Jill Callahan, Jessica Kostick, Cheyanne Warren, DeLacy Rhodes, and Dr. Jamie Sturgill, who have been a devoted support system for me throughout our time together at VCU. To have been able to share both the joy of success and the disappointment of failure with this group of individuals is a highlight of my past five years.

To my family and biggest fans, Doug, Cathy and Leslie Sale, I am deeply grateful for the strength, courage, and undying support they have provided me. They have each inspired me to dream big and persevere. Without their love and sustained encouragement, this journey, among all others, would not have been possible. I would like to extend a special thanks to my husband,

Brandon, for being my “rock”, and providing me with unwavering support on a daily basis. He is my best friend, my voice of reason, and challenges me to be my best in all things. I am incredibly fortunate to have each of these individuals in my life.

Table of Contents

Page

List of Tables

Page

List of Figures

Page

Abbreviations

°C	**degrees Celsius**
µL	**microliter**
µM	**micromolar**
5-aza-C	**5-aza-cytidine**
5-aza-dC	**5-aza-2'-deoxycytidine**
5hmC	**5-hydroxymethylcytosine**
5mC	**5-methylcytosine**
ATAD3A	**AAA-domain ATPase/protease 3A**
ATP6	**ATPase subunit 6**
AU	**absorbance units**
AUF1	**AU-rich element/poly(U)-binding/degradation factor 1**
BAH	**bromo-associated homology domain**
bp	**base pair**
CBP	**calmodulin binding peptide**
COX1	**cytochrome *c* oxidase subunit 1**
COX2	**cytochrome *c* oxidase subunit 2**
D-loop	**displacement loop**
DMAP1	**DNA methyltransferase associated protein 1**
DNMT1	**DNA methyltransferase 1**

DNMT3a	**DNA methyltransferase 3a**
DNMT3b	**DNA methyltransferase 3b**
ES cell	**embryonic stem cell**
EST	**expressed sequence tag**
EtBr	**ethidium bromide**
FBS	**fetal bovine serum**
gDNA	**genomic DNA**
GFP	**green fluorescent protein**
H3K4me3	**histone H3 lysine 4 trimethyl**
HCT116	**human colon carcinoma cell line**
HDAC	**histone deacetylase**
HMG	**high-mobility group protein**
Hsp60	**heat-shock protein 60kDa**
HSP1	**heavy strand promoter 1**
HSP2	**heavy strand promoter 2**
H-strand	**heavy strand**
IgG	**immunoglobulin**
LHA	**left homology arm**
LSP	**light strand promoter**
L-strand	**light strand**
MBD	**methyl-binding domain protein**
MeCP2	**methyl-CpG binding protein 2**
MeDIP	**methyl-DNA immunoprecipitation**

MEF	**murine embryonic fibroblast cell line**
mL	**milliliter**
MPP	**mitochondrial processing peptidase**
mtDNA	**mitochondrial DNA**
mTERF1	**mitochondrial transcription termination factor 1**
mtIP	**mitochondrial DNA immunoprecipitation**
mtRNAP	**mitochondrial RNA polymerase**
MTS	**mitochondrial targeting sequence**
mtTFB1	**mitochondrial transcription factor B1**
mtTFB2	**mitochondrial transcription factor B2**
NC1	**non-coding region 1**
NC2	**non-coding region 2**
ND1	**NADH dehydrogenase subunit 1**
ND6	**NADH dehydrogenase subunit 6**
ng	**nanogram**
nM	**nanomolar**
NRF1	**nuclear respiratory factor 1**
O_H	**origin of replication of the heavy strand**
O_L	**origin of replication of the light strand**
PAM	**presequence translocase-associated motor**
PCNA	**proliferating cell nuclear antigen**
PCR	**polymerase chain reaction**
PGC1α	**peroxisome proliferator-activated receptor γ-coactivator 1α**

POLG	**mitochondrial DNA polymerase gamma**
POLRMT	**mitochondrial RNA polymerase**
ProtA	**Protein A**
PTM	**post-translational modification**
qPCR	**quantitative polymerase chain reaction**
rAAV	**recombinant adeno-associated virus**
RFTS	**replication foci targeting sequence**
RHA	**right homology arm**
RT	**reverse transcriptase**
RT-qPCR	**reverse transcriptase quantitative polymerase chain reaction**
SAM	**S-adenosyl-L-methionine**
SDS-PAGE	**sodium dodecyl sulfide polyacrylamide gel electrophoresis**
TAP	**tandem affinity purification**
TCA	**tri-chloro-acetic acid**
TET1	**ten:eleven translocation protein 1**
TEV	**tobacco etch viral protease**
TFAM	**mitochondrial transcription factor A**
TFAM	**transcription factor of activated mitochondria**
TIM	**translocase of the inner mitochondrial membrane**
TOM	**translocase of the outer mitochondrial membrane**
UHRF1	**ubiquitin-like with PHD and ring finger domains 1**
VDAC	**voltage dependent anion channel**
WT	**wild-type**

β-gt	**β-glucosyltransferase**

ABSTRACT

FUNCTIONAL CONSEQUENCES OF CYTOSINE METHYLATION IN MITOCHONDRIAL DNA CATALYZED BY DNA METHYLTRANSFERASE 1

Lisa Sale Shock, B.A.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2011

Major Advisor: Shirley M. Taylor, PhD
Associate Professor, Department of Microbiology and Immunology

Cytosine methylation of mitochondrial DNA (mtDNA) was first described several decades ago, but neither the mechanism generating this modification nor its functional significance was known. Because mitochondrial dysfunction is a hallmark characteristic of numerous human diseases, including neurological and cardiovascular disease, aging and cancer, this dissertation addressed whether epigenetic modification of mtDNA regulates mitochondrial function. We show that mtDNA contains not only 5-methylcytosine (5mC), but also 5-hydroxymethylcytosine (5hmC), suggesting that previous reports likely underestimated the degree of epigenetic modification within the mitochondrial genome. We questioned how these modifications were generated by looking for mitochondrial isoforms of the nuclear-encoded DNA methyltransferases. We found that an isoform of the most abundant mammalian methyltransferase, DNA methyltransferase 1 (DNMT1) translocates to mitochondria, driven by

an in-frame mitochondrial targeting sequence (MTS) located upstream of the nuclear DNMT1 translational start site. This MTS is highly conserved across mammalian species, and directs a heterologous protein to the mitochondria. To investigate the function of mitochondrial DNMT1 (mtDNMT1), we created a cell line that carries a tandem-affinity purification (TAP) tag at the C-terminus of a single endogenous human DNMT1 allele. Using the DNMT1-TAP cell line, we showed that mtDNMT1 specifically binds mtDNA in a manner that is proportional to CpG density, proving its presence in the mitochondrial matrix. mtDNMT1 exhibits CpG-specific methyltransferase activity *in vitro* that is resistant to trypsin-treatment of intact mitochondria, but moderately susceptible to pharmacologic inhibition by the nucleoside analog 5-aza-2'-deoxycytidine (5-aza-dC). NRF1 and PGC1α, transcription factors that activate nuclear-encoded mitochondrial proteins in response to oxidative stress, were observed to up-regulate expression of mtDNMT1. Loss of p53, a tumor suppressor gene known to help control mitochondrial metabolism, also results in a striking increase in mtDNMT1 expression, and this up-regulation of mtDNMT1 appears to modify mitochondrial transcription in a gene-specific fashion. Our data suggests roles for mtDNMT1 in both the establishment and maintenance of cytosine methylation (from which 5hmC is presumably derived) and in the regulation of mitochondrial transcription. We propose that the enzymes responsible for epigenetic modification of mtDNA have potential as therapeutic targets, with relevance to a broad spectrum of human disorders.

Chapter 1: Introduction and Overview

I. Overview of mitochondrial structure and function

Mitochondria are considered to be the "powerhouse of the cell" for their essential role in electron transport and oxidative phoshorylation, which provide the majority of cellular energy (1, 2). These unique organelles, however, are more than just a powerhouse (2). Mitochondria are also involved in a number of other key cellular processes, including cell signaling, differentiation, metabolism, apoptosis, control of the cell cycle, and control of cell growth (1-3). Mitochondrial dysfunction contributes to the progression of various human pathologies, including neurodegenerative disorders, cardiac disease, aging and cancer (1-3). The involvement of mitochondria in such a wide range of diseases has made them an attractive therapeutic target among researchers spanning numerous fields.

Early studies on mitochondrial function used isolated systems, and focused on the organelle as an essential, yet autonomous, element of the cell. However, more recent findings strongly favor the view that mitochondria function as a cooperative, integrated system, highly involved in crosstalk with the nucleus and other components of the cell (1-4). Evidence supporting this crosstalk is based on the observation that a majority of the genes required for all of the mitochondria's biological functions are encoded by the nuclear genome (1-4).

Structurally, the mitochondrion is rather complex and unusual. It is comprised of two membranes that separate four distinct compartments, each with its own subset of functions (Figure 1-1). The outer mitochondrial membrane is composed of half lipid, half protein (5). It

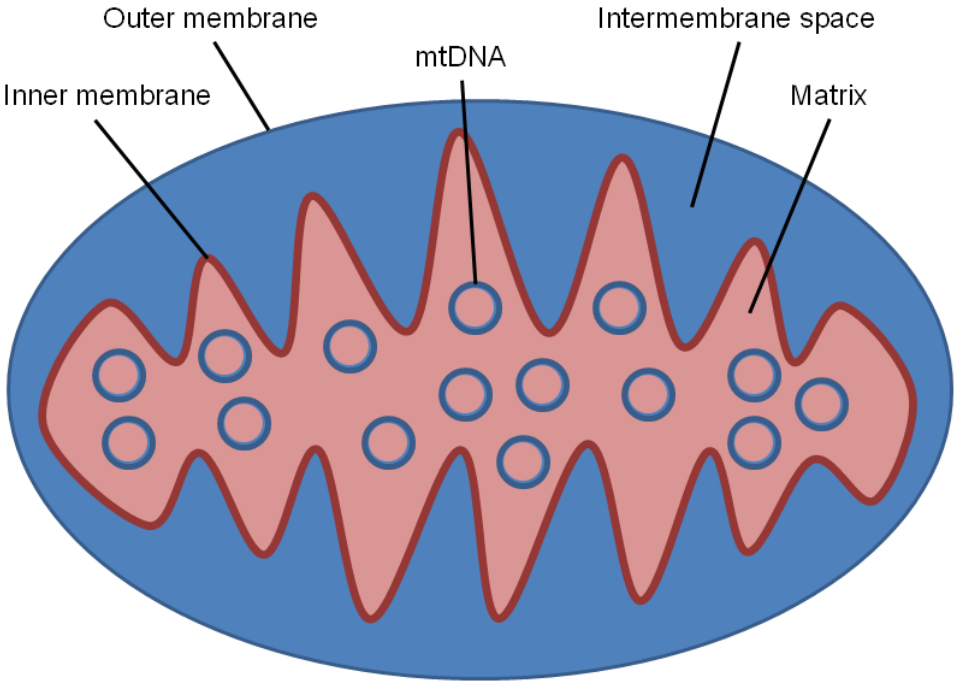


Figure 1-1: Cartoon of the organellar structure of a mitochondrion. Mitochondria are composed of two membranes (the outer and the inner membrane), which separate two aqueous compartments: the intermembrane space and the matrix. The matrix contains multiple copies of mitochondrial DNA (mtDNA). This compartmentalization assists the organelle in subdividing its many functions.

serves as the outermost barrier of the organelle, with a topology dotted by transmembrane channels, making it permeable to molecules less than 10 kDa in weight (5). The inner mitochondrial membrane is only about 20% lipid, 80% protein, making it a much denser barrier. The inner membrane's characteristic impermeability allows the formation of a proton-motive force, which drives ATP synthesis. Separating the outer and inner membranes is the aqueous intermembrane space. Its protein components are 100% nuclear in origin, and many contain N-terminal targeting signals that direct their import and translocation into mitochondria (6). The matrix is also an aqueous compartment, housing mitochondrial DNA (mtDNA) and ribosomes, and serves as the site for mitochondrial transcription, translation, and ATP synthesis.

Import of nuclear-encoded precursor proteins into mitochondria is most frequently driven by N-terminal targeting sequences that are recognized by receptors on the outer mitochondrial surface. Once imported through the mitochondrial outer membrane, targeting sequences then serve as sorting signals, directing the precursor protein to its destination (6, 7). Either during translocation, or shortly thereafter, the targeting sequences are cleaved by the mitochondrial processing peptidase (MPP) to yield their final functional form (6, 7).

II. General features of mtDNA, including mtDNA methylation

Part of what makes mitochondria so unique is that these organelles contain their own genetic system and the means for translating proteins encoded on the mitochondrial genome. Mammalian mitochondrial DNA exists as a 16.5kb double-stranded circular molecule. Despite its small size, mtDNA represents ~1% of a cell's total genomic DNA, because mtDNA is present in two to ten copies per mitochondrion, and 10^3-10^4 copies per cell (Figure 1-2) (1, 4).

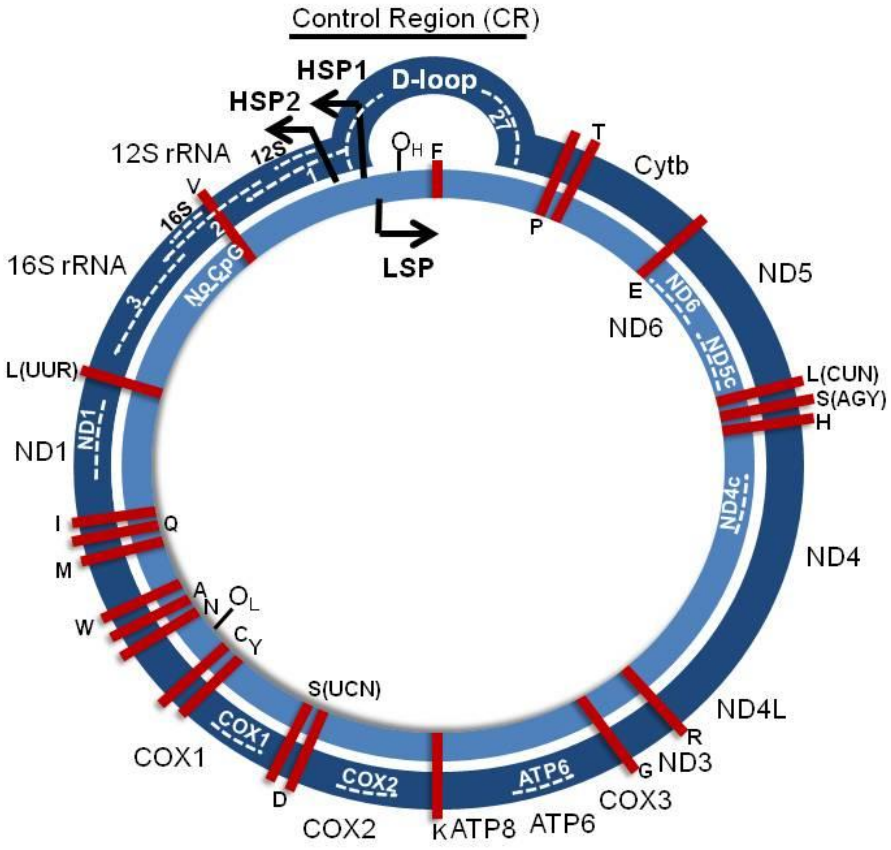


Figure 1-2: Schematic diagram highlighting general features of mtDNA. tRNA genes (red bars, single-letter code) intersperse the protein-coding and rRNA genes on both the H-strand (dark blue) and L-strand (light blue). Positions of amplicons used in MeDIP and mitochondrial gene expression studies (Chapter 2) are represented by dashed white lines. HSP1, heavy strand promoter 1; HSP2, heavy strand promoter 2; LSP, light strand promoter.

The mitochondrial genome exhibits an extreme economy in its organization: it is relatively small in size, and contains no introns or non-coding regions. Both strands of mitochondrial DNA are used to encode a total of 13 proteins, present in the respiratory chain complexes, 2 rRNAs and 22 tRNAs (Figure 1-2), by far an insufficient collection of proteins to allow for complete functioning of the organelle. As mentioned above, the vast majority of "mitochondrial" proteins, including those required for mtDNA replication and transcription, are encoded in the nucleus and transported into the mitochondria.

In contrast to the nuclear genome, mtDNA is not complexed with histones. However, mtDNA is present in stable protein-containing complexes called nucleoids, each containing multiple copies of mtDNA bound to a complex mixture of proteins, including the mtDNA polymerase (POLG), Twinkle helicase and various transcription factors (8, 9). These mtDNA nucleoids are thought to function in a manner similar to nuclear histones by packaging between five and seven copies of mtDNA each into a restricted diameter of only 70nm (9). This compaction property of nucleoids therefore contributes to mtDNA segregation and heritable transmission of mitochondrial genomes into daughter mitochondria during fusion and fission (8-10).

Mitochondrial nucleoid complexes are proposed to exist in two zones: 1) the core complex, which contains mtDNA actively undergoing replication and/or transcription, associated with appropriate POLG and mtRNA polymerase (mtRNAP) proteins, and 2) the peripheral complex, containing mtDNA and natively-bound proteins that are not efficiently crosslinked to the DNA, including heat-shock protein 60kDa (Hsp60) and AAA-domain ATPase/protease 3A (ATAD3A) (9). Among the most abundant components of mtDNA nucleoid core complexes is mitochondrial transcription factor A (TFAM) (8, 9). A number of recent studies have reported

levels of TFAM in sufficient quantities to coat the entire mitochondrial genome, supporting its proposed role in binding, bending and compacting mtDNA into distinct foci (10, 11), and protecting mtDNA from oxidative damage (12), in addition to its role as a mitochondrial transcription factor (10).

Isolation of mitochondrial replicative intermediates from mouse L cells using dye-CsCl gradients and electron microscopy suggested a model in which mtDNA replication originates at the origin of replication of the heavy strand (O_H), upon formation of a displacement loop (D-loop) (4). This D-loop exists as a triplex structure containing a short daughter heavy (H)-strand RNA, which remains stably associated with the parental strand, serving as a primer to initiate H-strand synthesis (4). Once initiated, synthesis of the H-strand proceeds unidirectionally for the entire length of the molecule. Upon completing approximately three-fourths of the H-strand, synthesis of the L-strand begins at its origin (O_L), and continues in the opposite direction. Once replication has traversed the full length of the molecule in both directions, both H- and L-daughter strands are converted to covalently-closed circles and super-helical turns are added, forming traditional super-helical mtDNA (4).

Studies performed decades ago first reported the presence of 5-methylcytosine (5mC)-modified residues in mtDNA. These early findings were controversial because of technical limitations and it was considered likely that the modifications measured were reflective of contaminating nuclear DNA. Using methylation-sensitive restriction endonuclease cleavage and nearest-neighbor analysis, these groups reported low levels of methylation restricted to CpG dinucleotides (only 3-5% of CpGs were methylated), in the mitochondria of several species (15-17). This methylation was described as “nonrandom”, meaning that, while certain sites were

most frequently modified, different sites were found to be methylated to a different extent (16, 17).

Interestingly, methylation is not the only feature mtDNA shares with nuclear DNA. Mitochondrial DNA was also found to exhibit an underrepresentation of CpG dinucleotides (termed "CpG suppression") to nearly the same extent as the nuclear genome (18). The phrase "CpG suppression" refers to the general rarity of CpG dinucleotides within DNA, compared to the random expected frequency. 5mC, which occurs in the mammalian genome almost exclusively at CpG dinucleotides, is a known mutational "hotspot", as it can be spontaneously deaminated to thymidine. Thus, CpG suppression is thought to be reflective of a functional significance of 5mC residues and lends further support to the idea that mtDNA was at some time, or is now, methylated.

Although the presence of mtDNA methylation has been known for quite some time, it has remained surprisingly underexplored. Thus, the remainder of the overview on epigenetic mechanisms and their roles in regulating gene expression will focus on what is known about these processes in the nuclear compartment, and consider whether these mechanisms might also operate in mitochondria.

III. Epigenetic regulation of gene expression in the nucleus

There is a clear and well-established role for genetics in human disease, but nuclear genetic information is packaged in a complex chromatin structure and the information content extends beyond the sequences of genes (Figure 1-3A). The term epigenetics refers to the extra layer of instructions that influences gene activity without altering the DNA sequence. Epigenetic modifications exist in three main forms: 1) DNA methylation, 2) histone modifications, and 3)

nucleosome positioning (Figure 1-3B). Mechanistically, each of these mechanisms induces dynamic and dramatic changes in chromatin structure, thereby dictating gene expression and inheritance (19). Epigenetic modifications have fundamental roles in normal development and biology, and are often among the first factors to become dysregulated in disease. Although epigenetic modifications are somatically heritable, they are also reversible, making them the subject of intense investigation as therapeutic targets (20, 21). Aberrations to the epigenetic landscape have been demonstrated in cancer, diabetes, aging, and a number of neurological disorders (21).

While the human genome is essentially the same in all of our cells, the epigenome differs in a tissue-specific fashion, and can be influenced by a cell's environmental cues. Thus, epigenetics provides a platform for the generation of multiple different phenotypes arising from the same genotype (22). The tissue-specificity of epigenetic marks contributes to cellular differentiation and, when aberrantly placed, can lead to neoplastic transformation with characteristics specific to the host organ (22, 23). Epigenetic mechanisms also have a "nurture" component, such that changes to cellular environmental conditions, including physical, chemical or dietary influences, can induce a stable, heritable epigenomic response (24). This gene-

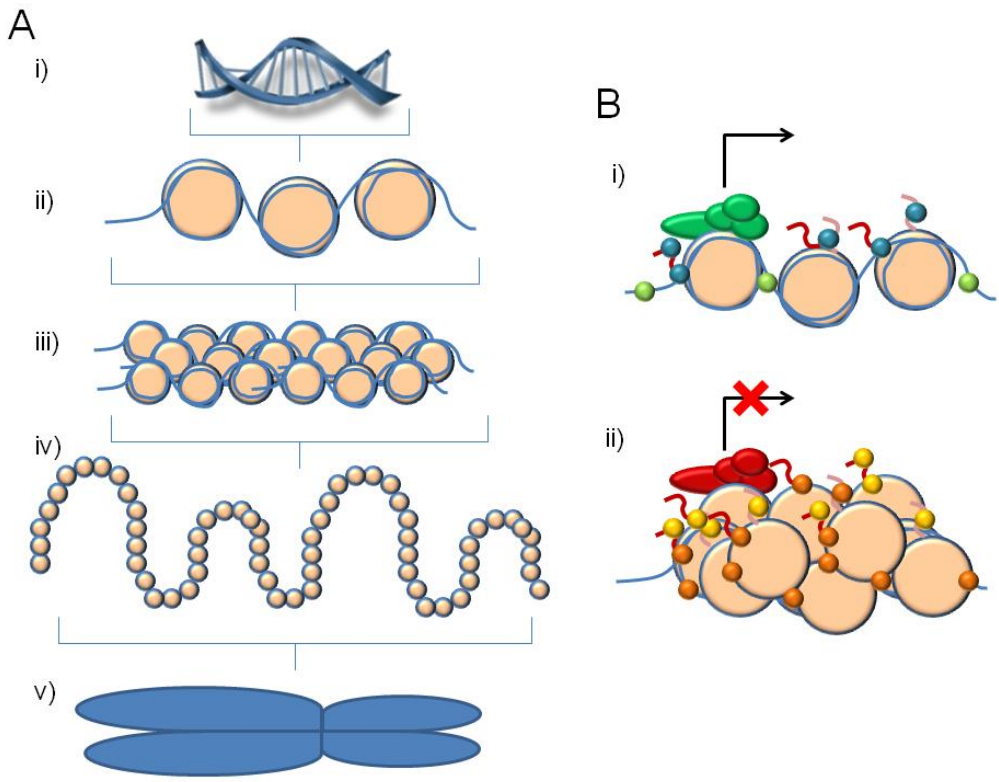


Figure 1-3: Structural organization and packaging of nuclear genetic material. (A) The organization of DNA within the chromatin structure: (i) A short region of DNA double helix; (ii) DNA is wound around a histone octamer, forming nucleosomes separated by short stretches of linker DNA; (iii) nucleosomes condense to form a packed chromatin fiber; (iv) chromatin fibers are further folded into higher-order structures; (v) the full mitotic chromosome. (B) Euchromatic and heterochromatic states of chromatin. (i) Euchromatin, represented by a transcriptionally active gene with an open chromatin structure, sparse DNA methylation (green circles), and activating histone marks (blue circles). The transcriptional machinery is shown present at the promoter, represented by a complex of green ovals. (ii) Heterochromatin, represented by a transcriptionally silenced gene with closed chromatin structure, dense DNA methylation (orange circles), and repressive histone marks (yellow circles). Adapted from figures by Felsenfeld G, Groudine M (2003) *Nature* and Feinberg AP (2007) *Nature*.

environment interaction in early development is thought to be linked to disease outcomes later in life (24).

DNA methylation is probably the most well-studied of epigenetic mechanisms. DNA methylation is the result of a stable chemical alteration of the DNA, and can have long-term consequences on gene activity. It was first suggested by Riggs (25) and Holliday and Pugh (26) that patterns of DNA methylation vary in different cell types, and that these methylation patterns are capable of altering gene expression profiles. It was hypothesized that the mechanisms underlying these effects on gene expression acted through modulation of DNA binding by transcription factors or other proteins (25, 26). Molecular studies have confirmed these initial hypotheses, demonstrating that the covalent addition of a methyl group to the DNA changes the facade of the major groove, altering accessibility by DNA binding proteins (27, 28). Although DNA methylation is generally associated with gene silencing, in actual fact, depending on whether a positive or negative regulatory element is affected, DNA methylation can act to either increase or decrease transcription (27). The precise function of methylation depends on the sequence context surrounding the modified residue, and its proximity relative to the transcriptional start site of the adjacent gene (27).

Post-translational modifications (PTMs) of histone proteins, collectively, comprise another mechanism influencing the recruitment of effector protein complexes that regulate transcription (28). Histone modifications include acetylation, methylation, phosphorylation, SUMOylation and ubiquitination (29). The majority of these modifications decorate either the N- or C-terminal histone tails; however, some localize to the globular domains of the histone proteins (28). Whereas both acetylation and methylation involve the addition of small chemical groups, SUMOylation and ubiquitination add larger, more bulky groups that can presumably

have a more profound impact on chromatin structure. It was first thought that histone PTMs could be classified by their consistent behavior as either positive or negative marks, however, the functions of individual histone modifications are neither easily definable nor predictable. Rather, like DNA methylation, the consequences of each mark are highly dependent on the specific residues and contexts in which they are added (28).

The coordinated effects of DNA methylation and histone modification result in an "epigenetic code," a mosaic pattern of factors regulating gene expression at the transcriptional level. The combination of modifications can either work directly, by influencing local chromatin structure, or indirectly, through the recruitment of effector molecules (28). Chromosomes have two structurally and functionally distinct chromatin topographies, distinguishable by the positioning of nucleosomes along the chromatin fiber. The term "euchromatin" has been assigned to the active transcriptional state, whereby chromatin is dotted by acetylated histones and devoid of methylation (Figure 1-3B). This open configuration allows for the recruitment of activating factors, and accessibility to DNA by the transcriptional machinery. In contrast, the term "heterochromatin" refers to the transcriptionally inert state of chromatin that adopts a closed, compact configuration, preventing access by the transcriptional machinery. Heterochromatin is characterized by an abundance of DNA cytosine methylation, and hypoacetylation of histones. The prevailing thought is that these epigenetic marks function cooperatively, to either activate gene expression or to establish repression (28); in this way, the dynamic state of chromatin is regulated by two opposing sets of epigenetic marks for fine-tuned control.

IV. Dynamics and mechanisms of DNA methylation and demethylation

DNA methylation arises through the covalent addition of a methyl group to the 5' position of a cytosine, and occurs almost exclusively at CpG dinucleotides within the mammalian genome. CpG dinucleotides are under-represented in mammalian DNA, due in part to the mutagenic pressure of spontaneous 5-methylcytosine (5mC) deamination, yet vertebrate genomes contain the highest levels of 5mC modifications in the animal kingdom (30). In humans, approximately 70-80% of all CpG dinucleotides are methylated, meaning that 5mC accounts for approximately 1% of all DNA bases (31).

Although they represent long-term heritable marks, DNA methylation patterns exhibit striking temporal and spatial variation. During early development, the paternal pronucleus undergoes a distinct phase of demethylation, which erases the existing methylation pattern and essentially reprograms the genome (30, 32). The somatic DNA methylation pattern is established by a wave of *de novo* methylation just before implantation, and that pattern is then propagated during DNA replication throughout development of the organism (30, 33). DNA methylation patterns in somatic cells and the adult organism are stable, except for a distinct subset of genes that undergo selective demethylation when their expression is needed, as in the demethylation of imprinted genes in primordial germ cells (32). It is thought that DNA methylation patterns are maintained between cell generations by a mechanism based on semi-conservative replication, whereby the parental pattern is copied onto the nascent DNA strand (34).

Mammals exhibit a distinct bimodal pattern of methylation, with hypermethylation of CpG dinucleotides sparsely distributed throughout the genome, and hypomethylation of "CpG-islands," which are short stretches of CG-dense sequence surrounding the promoter regions of genes (Figure 1-4). Methylated CpGs throughout the body of a gene are thought to contribute to

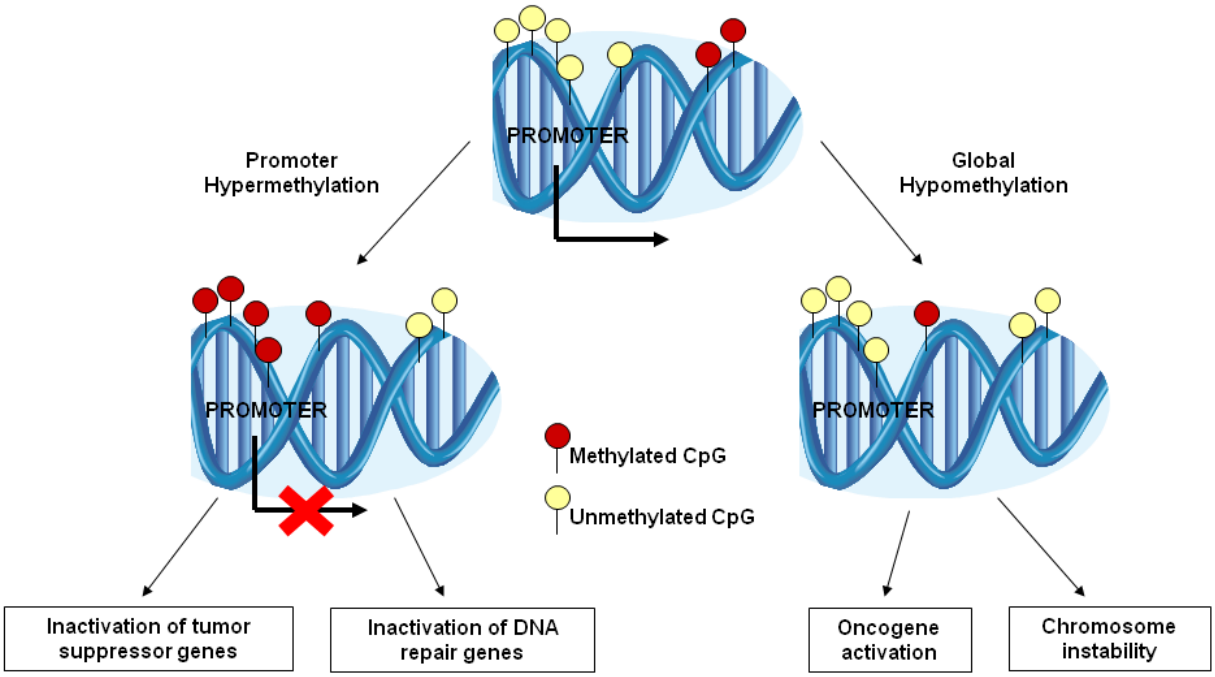


Figure 1-4: Roles of DNA methylation in normal development and disease states. DNMT1 functions normally to faithfully maintain the established pattern of methylation throughout DNA replication. When functioning abnormally, as in cases of over-expression, DNMT1 has been shown to methylate DNA *de novo*, and appears to drive tumor progression. The mechanisms forcing these dramatics shifts in DNA methylation remain to be understood, but it is apparent that the factors involved would be extremely relevant therapeutic targets. Adapted from figure by Taylor, S.M. (2006) *Cell Science Reviews*.

genomic stability, whereas the absence of methylation at CpG-island-associated promoters supports transcription and expression of the adjacent genes. Cells undergoing transformation to a neoplastic phenotype exhibit the opposite methylation profile: global hypomethylation, leading to genomic instability and hypermethylation of CpG-island-associated promoters, resulting in transcriptional silencing (22, 23).

The enzymatic reaction of DNA methylation is catalyzed by a family of DNA (cytosine-5) methyltransferases (DNMTs), whose individual roles will be discussed in more detail below. These enzymes require the aid of cofactor and methyl group donor, S-adenosyl-L-methionine (SAM) and rely on the steric relief of base-flipping the target cytosine out of the DNA backbone (35, 36). The DNMTs exert a nucleophilic attack on the carbon-6 position of the target cytosine. This attack is performed by a thiol group positioned within the ProCys active site of the enzyme, and results in the formation of a covalent complex between the DNMT and the cytosine C-6 position. Formation of the covalent intermediate activates the carbon-5 position of the target cytosine for electrophilic attack, which leads to the transfer of a methyl group from SAM to the C-5 position of the cytosine (35). The covalent reaction intermediate is resolved when the 5-6 double bond is reformed and the DNMT is released by β-elimination (36). The active site cysteine residue has been shown to be essential for both covalent complex formation with the target cytosine residue, and catalysis of the DNMT enzyme (37, 38).

The requirement for DNA methylation in normal mammalian development is demonstrated by a number of key studies which showed that pharmacological inhibition of DNA methylation by nucleoside analogs of cytidine (5-aza-cytidine and 5-aza-deoxycytidine) induced striking phenotypic changes in cultured mouse cells (39, 40). Multinucleated myotubes, a characteristic of contractile striated muscle cells, adipocytes and chondrocytes were all observed

to arise from cultures of 10T1/2 and NIH/3T3 mouse cells treated with the cytidine analogs (39). The resultant changes in the differentiated state of treated cells were traced back to changes in the methylation patterns of their DNA (40). Subsequent deletion of the genes encoding the DNMT enzymes resulted in an approximate 3-fold reduction in genomic 5mC levels (41). While ES cells carrying the deletions could survive in culture without obvious defects in viability or proliferation, knockout mice exhibited severe developmental abnormalities, and died early in development (41, 42).

Just as there are mechanisms in place to add methyl groups to DNA, the process of DNA demethylation also has been examined at both the genome-wide and locus-specific levels (43). However, the precise mechanism underlying removal of 5mC has been under intense debate. For years researchers sought an active demethylase, and several reports claimed to have identified known proteins with secondary demethylating functions (43). Many hypothesized that such a demethylation reaction could occur through a base excision repair-like process, in which the 5mC base is excised and replaced with an unmethylated cytosine (43, 44). Active DNA demethylation has been well-characterized in plants, and involves a group of 5mC-specific DNA glycosylases that initiate base excision repair of 5mC to unmethylated cytosines (44).

The recent discovery of 5-hydroxymethylcytosine (5hmC) residues in mammalian DNA sparked speculations that 5hmC could represent an intermediate in either active or passive demethylation of mammalian DNA (45, 46). Earlier this year, a study described a two-step mechanism mediated by the ten-eleven translocase (TET) family of methylcytosine oxygenases, whereby 5mC residues are first converted to 5hmC residues through oxidation by the TET1 protein. 5hmC residues are then deaminated by the activation-induced deaminase

(AID)/apolipoprotein B mRNA-editing enzyme complex (APOBEC) family of deaminases to form uracil, which is repaired by the base excision repair pathway (47).

The predominant mechanism for generation of 5hmC modification is thought to be through oxidation of 5mC residues by the TET1 protein, although other mechanisms have been suggested, including the direct addition of an aldehyde to DNA by DNMT1 (Figure 2-15) (46-48). It is curious that, through two alternate mechanisms, DNMT1 could be responsible both for generating 5mC modifications and removing them by initiating demethylation, but this level of mechanistic fine-tuning has been described for other enzymes participating in "suicide" inhibition of their own activity, such as acetylcholinesterase (49).

5hmC was first identified in brain and liver tissues extracted from several species, and estimated by thin layer chromatography to represent approximately 15% of the total number of cytosines within the rat genome (50). This is far greater than the estimated frequency of 3-6% for 5mC residues among cytosine bases within the normal mammalian genome (51), but is likely either an artifact, or a reflection of the tissue-specificity of the 5hmC modification, as the highest levels of 5hmC are undoubtedly found in the brain (52). More recent enzymatic quantitations of 5hmC levels in genomic DNA from murine embryonic stem cells estimate that the modified base occurs at around 0.3% of cytosines (53). Also unlike 5mC, 5hmC is enriched at gene promoters and CpG islands, and co-localizes with euchromatin, which promotes high levels of transcription (54). Thus, the distribution of 5hmC modifications also demonstrates a bimodal pattern, but opposite to that for 5mC, which further supports the idea that 5hmC is derived from pre-existing 5mC modifications (46, 54).

Just as the presence of 5mC can have either direct or indirect effects on transcription, 5hmC is also thought to influence local chromatin structure by two potential means: 1) recruiting

5hmC-specific proteins or 2) preventing binding by 5mC-binding proteins (52). Studies have shown that some 5mC-specific binding proteins, including methyl-CpG binding protein 2 (MeCP2), and methyl-binding domain proteins, MBD1, MBD2 and MBD4, do not recognize 5hmC residues (55, 56). Additionally, certain restriction endonucleases, including the *Msp*I and *Hpa*II isoschizomer can distinguish between 5mC and 5hmC modifications at the internal cytosine within their 5'-CCGG-3' target restriction site when first modified by β-glucosyltransferase (52). However, methods for specific detection and quantitation of 5hmC are still lacking; the vast majority of methods commonly employed to map and measure DNA methylation, such as classical sodium bisulfite sequencing, cannot distinguish between 5mC and 5hmC modifications (53, 56).

V. Roles of the DNA methyltransferase (DNMT) enzymes

While several members of the family of DNA methyltransferase homologs are known to exist (Figure 1-5), only three have DNA methyltransferase activity: DNMT1, DNMT3a and DNMT3b. The *de novo* methyltransferases, DNA methyltransferase 3a (DNMT3a) and DNA methyltransferase 3b (DNMT3b), are thought to establish the pattern of methylation on unmethylated DNA during early embryogenesis (34). That pattern is then faithfully maintained throughout subsequent rounds of DNA replication by the most abundant mammalian methyltransferase, DNA methyltransferase 1 (DNMT1) (34). DNMT2 is a tRNA methyltransferase, and DNMT3L does not have any catalytic activity, but is thought to function as a regulatory factor, interacting with and stimulating methylation by DNMT3a (34, 57).

All DNA (cytosine-5) methyltransferases share similarities in sequence and structure. All (except for DNMT2) consist of two major domains: the N-terminal regulatory domain, and the

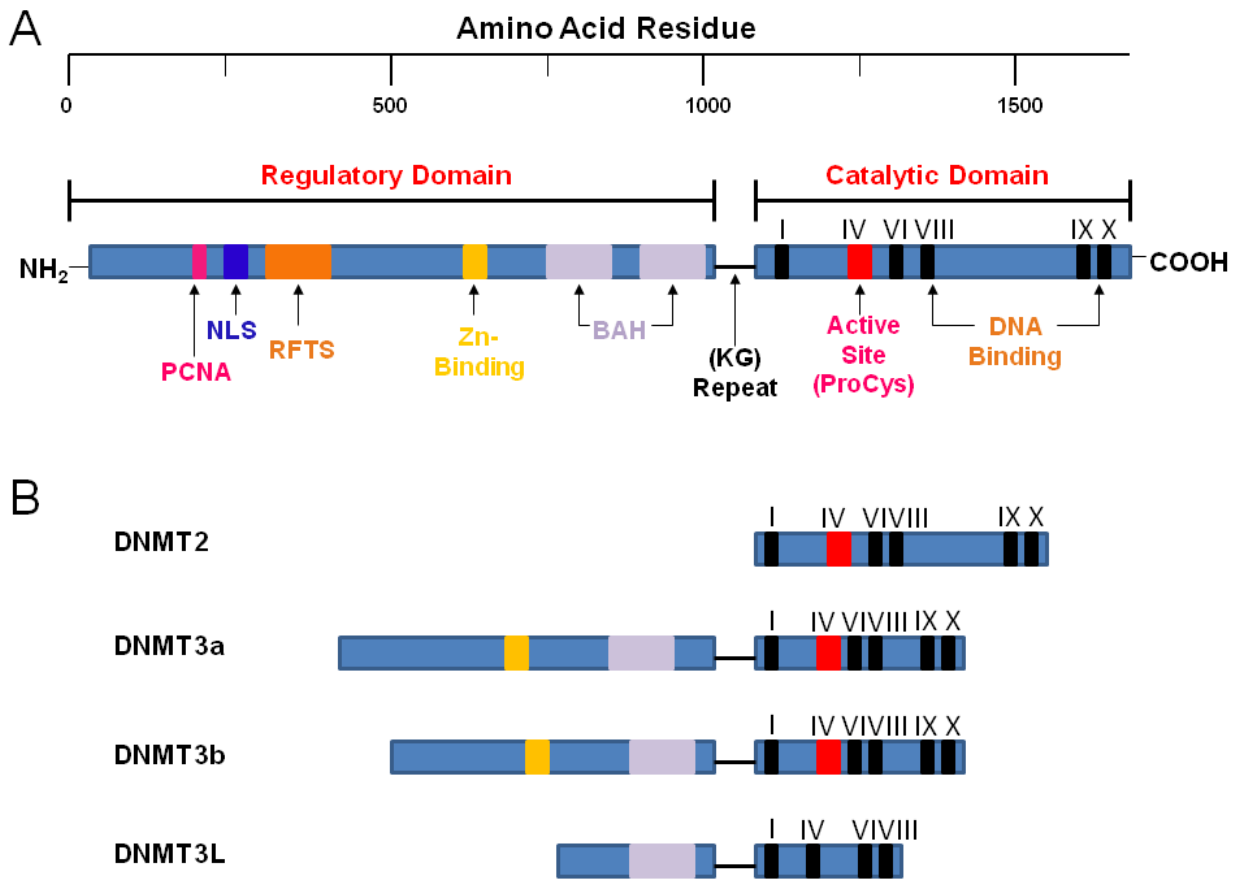


Figure 1-5: Basic molecular structures of for the family of DNMT enzymes. (A) The structure of DNMT1 consists of two domains: the N-terminal regulatory domain and the smaller C-terminal catalytic domain. The regulatory domain contains binding sites for PCNA (PCNA) and other proteins, a nuclear localization signal (NLS), a replication foci targeting sequence (RFTS), CXXC Zn-finger domains (Zn-binding) and bromo-associated homology (BAH) domains. The catalytic domain contains conserved motifs (black bars labeled with roman numerals), each with specific functions contributing to enzymatic activity. Motif IV is the active site; motifs VIII and X bind the target cytosine residue. The N-terminal and C-terminal domains are separated by a LysGly (KG repeat) hinge region. (B) Organization of structural domains and conserved motifs among the DNMT enzymes. The C-terminal domains for all other DNA (cytosine-5) methyltransferases share strong amino acid sequence homology and conserved motifs. The N-terminal domains are highly variable across the DNMTs. Adapted from figure by Hermann A, Gowher H, Jeltsch A (2004) *Cell Mol Life Sci*.

C-terminal catalytic domain. Like bacterial DNMTs, DNMT2 lacks the large N-terminal domain (Figure 1-5). The regulatory domain serves to facilitate interactions with numerous proteins and binds to DNA (57). The catalytic domain, on the other hand, serves as the active center of the enzyme, and houses the conserved motifs characteristic of all eukaryotic and prokaryotic cytosine-5 DNMTs. Motifs I, II and III form a binding pocket for SAM; motif IV contains the ProCys dipeptide initiating the methyl transfer reaction; motifs V, VII and IX are involved in preserving the structure of the target recognition domain; motifs VI, VIII and X form the binding pocket for the cytosine substrate (Figure 1-5) (36, 57).

The initial hypothesis that *de novo* and maintenance methylation were two distinct processes arose from observations by Lei, *et al.* (58). They used homologous recombination to delete the only known DNA methyltransferase gene at the time (now known as DNMT1), and to their surprise, the DNMT-null cells retained low but stable levels of 5mC and DNMT activity. These results suggested that not only was an independently-encoded DNMT present in mammalian cells, but it contained enzymatic activity capable of methylating exogenously-introduced unmethylated DNA *in vivo* (58). Upon closer examination of the EST database, a total of four human EST clones were found to contain the highly conserved DNA (cytosine-5) methyltransferase motifs (59). Two of these clones shared a high degree of sequence similarity, but differed from DNMT1 and DNMT2; these were subsequently named DNMT3a and DNMT3b, although they are encoded by separate genes (59). Both DNMT3a and DNMT3b were found to be highly expressed in undifferentiated murine embryonic stem cells, but barely detectable in differentiated embryoid bodies and adult tissues (59). DNMT1, on the other hand, is fairly ubiquitously expressed in mammalian cells, but has been shown to be indispensable for

proper differentiation and survival of differentiated cells (60). Thus, the expression patterns for these enzymes correlate with their proposed roles as *de novo* or maintenance methyltransferases.

Additional genetic studies confirmed that DNMT3a and DNMT3b were responsible for establishment of the methylation profile (42). Embryonic stem (ES) cells carrying homozygous deletions in either DNMT3a or DNMT3b displayed no change in their ability to methylate unmethylated proviral DNA, however, targeted disruption of the active site in both genes (DNMT3a-/-, DNMT3b-/-) prevented *de novo* methylation activity (42). Similarly, mice carrying a double mutation for both enzymes exhibited a more severe phenotype than either individual deletion, suggesting an overlap in functions among the DNMTs (42).

As their names suggest, the catalytically active DNMTs work quite differently on DNA substrates. DNMT3a and DNMT3b both prefer unmethylated DNA, but DNMT3a modifies DNA in a distributive manner (61), while DNMT3b is processive in its activity (62). This difference in mechanism is thought to be related to the specific function for each enzyme. For instance, DNMT3a has been shown to establish methylation patterns at single copy genes, which requires recruitment for each methylation event (63), while DNMT3b methylates regions of high CG content within pericentromeric repeats, which is benefited by a processive mechanism (35). DNMT1 initially garnered the name "maintenance methyltransferase" because it strongly prefers hemi-methylated DNA substrates, although it is quite active on unmethylated DNA, with a faster catalytic rate, in fact, than either of the "*de novo*" DNMTs (64). DNMT1 has also been shown to co-localize with sites of active DNA replication and associate with the replication machinery, observations which both support its role of copying the parental methylation pattern onto the daughter DNA strand (57).

Due to the dynamic nature of DNA methylation and the apparent differences in mechanism and function of the DNMT enzymes, it has long been accepted that two distinct methylation processes (*de novo* versus maintenance methylation) are responsible for generation and establishment of the DNA methylation profile. However, recent literature proposes a new model for DNA methylation, suggesting that the roles of DNMTs may not be as mutually exclusive as previously thought (65). Among the first suggestions of cooperativity among the enzymes came from Bert Vogelstein's laboratory, which genetically disrupted first DNMT1 and then DNMT3b in a human colon cancer cell line (66). Interestingly, cells lacking either DNMT1 or DNMT3b alone retained a substantial level of genomic methylation, and methylation-induced gene silencing of repetitive DNA sequences, and imprinting and tumor suppressor genes (66). However, deletion of both DNMT1 and DNMT3b in double knockout (DKO) cells resulted in approximately 95% reduction in DNA methylation levels, and almost complete loss of methyltransferase activity (66). It should be noted that the strategy behind generation of these DNMT1 knockout (KO) cell lines involved targeted disruption of the N-terminus, replacing exons 3-5 with a hygromycin resistance cassette. It was later discovered that these DNMT1 "knock-outs", were actually hypomorphs, and contained alternatively-spliced, truncated versions of the DNMT1 protein that retained *in vitro* catalytic activity (67). Interestingly, however, among the sites that underwent complete demethylation in the DKO cells was the CpG-island promoter of *DNMT3a2*, which is a variant of DNMT3a (67), and its up-regulation in cells deplete of the other DNMTs may represent some sort of compensatory mechanism to maintain sufficient levels of methylation for cell survival.

Descriptions of cooperation between the DNMTs were substantiated by others who reported an interaction, nuclear co-localization, and some degree of communication between the

de novo enzymes, DNMT3a and DNMT3b, and the maintenance enzyme, DNMT1 as they facilitated methylation spreading (68, 69). Interestingly, recent reports described striking differences in binding affinities of the DNMTs to select regions of chromatin; DNMT1 interacted more with linker DNA, while DNMT3a and DNMT3b were both tightly anchored on nucleosomes (70). Although the DNMTs demonstrate this apparent distinction in their interaction with DNA at the chromatin level, it suggests that they may in fact be compensating, or "covering for" each other by methylating sites missed by the other enzyme(s) (70).

Collectively, a closer examination of the specific activities of each DNMT enzyme on different DNA substrates, their unique interactions with chromatin, and the results of genetic studies disrupting the DNMTs alone or in combination, have led researchers to conclude that it is the differences in the functional properties of these enzymes that drives their cooperation. Classifying the DNMTs as either all "*de novo*" or all "maintenance" is therefore an over-simplification. Certain aspects of each DNMT are required at different times throughout development, and along DNA of variable sequence contexts to achieve the same general end. These aspects are exploited by nature to achieve the best possible cooperative mechanism for ensuring proper maintenance of DNA methylation.

VI. DNA methyltransferase 1 (DNMT1) expression and regulation

Because DNA methylation exerts such fine-tuned control over gene expression, it should be expected that the cellular levels and activity of the major methyltransferase, DNMT1, are also tightly controlled. DNMT1 is the largest and most abundant of the mammalian DNMTs (Figure 1-5), with coding sequence totaling over 1600 amino acid residues, and producing a protein of approximately 185kDa in size, which is roughly three times larger than its prokaryotic

counterpart (57). The unique tripartite structure of DNMT1 is thought to be the result of a fusion of three distinct genes, one of which is an ancestral prokaryotic methyltransferase and believed to give rise to the C-terminal catalytic domain (71). The N-terminal domain itself exhibits a bipartite structure, with two distinct regions: one dispensable for enzymatic activity, the other essential, although they encode redundant nuclear localization signals (Figure 1-5, RFTS and NLS) (71).

Also within the N-terminal domain of DNMT1 is a replication focus targeting sequence (RFTS), located between amino acids 110-307, that directs the enzyme to sites of active DNA replication in the nucleus of the cell (Figure 1-5) (35, 72). The localization of DNMT1 is highly cell cycle-dependent, as its co-localization with replication foci occurs only during S-phase; during all other stages of the cell cycle, DNMT1 exhibits a diffuse nuclear distribution (72). Two bromo adjacent homology (BAH) domains are located within the N-terminal regulatory domain of DNMT1. These are commonly found in chromatin-associated proteins and have been proposed to serve as protein-protein binding sites within DNMT1 (35, 57). In close proximity to the RFTS and BAH are a number of other protein binding domains, including a proliferating cell nuclear antigen (PCNA) interacting domain, which also facilitates its recruitment to the DNA replication fork and sites of DNA repair (57). Histone deacetylase 1 and 2 (HDAC 1 and 2), DNA methyltransferase associated protein 1 (DMAP1) and retinoblastoma tumor suppressor (Rb) also all have binding sites within the N-terminal domain of DNMT1, which link processes of DNA methylation, histone deacetylation and a growth regulatory pathway (35, 57). Lastly, the center of the N-terminal domain of DNMT1 contains a cysteine-rich Zn^{2+} binding region, comprised of two CXXCXXC clusters with two isolated cysteines, that strongly resembles zinc-finger binding domains of other proteins (73). This region was examined for its DNA binding

properties and found to be essential for the methyltransferase activity of DNMT1, demonstrating that the N-terminal domain of the protein cooperates with the C-terminal, catalytic domain (73).

The C-terminal domain of DNMT1 is the active center of the enzyme, containing the active site that is absolutely required for enzymatic function (Figure 1-5). As mentioned above, the C-terminal domain is not catalytically active in isolation; it requires the presence of at least the majority of the N-terminal domain to assist in its function through an allosteric activation mechanism (57, 74).

Several amino acids residues of the DNMT1 protein have been shown to be altered by post-translational modifications, each with unique consequences dependent on where the modification occurs. For instance, Akt kinase-mediated phosphorylation of Ser209 within the nuclear localization signal of the N-terminal domain increases nuclear translocation of DNMT1 from the cytosol (75). Phosphorylation of Ser514 and Ser515 are thought to play roles in targeting of the enzyme to replication foci and allosteric activation of methyltransferase activity, respectively (76, 77). Although post-translational modifications are common events, modulating the activity, localization, stability and interactions of the proteins they modify, the PTMs for DNMT1 remain largely unexplored.

As briefly discussed previously, there is a cell cycle component to regulation of DNMT1 expression and activity. Evidence for this type of control stems from findings that cellular levels of DNMT1 transcript and protein are positively correlated with cell proliferation (78). DNMT1 expression was minimal in G_0/G_1 and peaked late in S-phase of synchronized mouse cells (78). The timing of these fluxes in DNMT1 levels aligned with levels of DNA synthesis, suggesting that DNA methylation and replication are co-regulated by the cell cycle.

A current subject of intense investigation is the regulation of DNMT1 function through binding by accessory proteins, with one of particular interest: ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1). UHRF1, also known as Np95 in mouse and ICBP90 in human, has been shown to be required for maintenance of DNA methylation patterns in mammalian cells, a result of its direct interaction with DNMT1 (79). UHRF1 contains a set and ring-associated (SRA) domain that exhibits strong hemi-methylated DNA binding properties. UHRF1 also binds DNMT1, and is therefore thought to play a role in tethering DNMT1 to sites of hemi-methylated CpGs, facilitating the faithful propagation of the established methylation pattern by DNMT1 (79). It has also recently been reported that, unlike MeCP2 or the methyl-binding domain proteins, UHRF1 recognizes and binds to 5hmC residues with a similar affinity to 5mC (80). It will be interesting to see if an interaction between UHRF1 and the TET family of proteins is revealed, and what role, if any, UHRF1 may play in DNA demethylation.

Aberrations to the DNA methylation profile have been considered early events in tumorigenesis (81). Numerous reports described elevated levels of DNMT1 in tumor cells, 4-3000-fold higher than in non-tumor cells, and the change in DNMT1 expression was shown to positively correlate with cancer stage progression (81, 82). Higher DNMT1 expression was also paired with abnormal methylation activity: greater overall genomic methylation and *de novo* methylation activity (83). However, upregulation of DNMT1 is not equivalent to dysregulation of DNMT1. Rather, aberrations to the DNA methylation pattern and distribution, not simply increases in methylation levels, are found in diseased states.

VII. Summary and Objectives

Our laboratory has intensely pursued the molecular mechanisms underlying the dysregulation of the most abundant mammalian DNA methyltransferase, DNMT1, as they relate to the aberrant DNA methylation patterns of diseased cells. A great deal is known about epigenetic mechanisms and their regulatory effects on transcription of the nuclear genome, but the field of mitochondrial epigenetics severely lacks exploration. The early findings that mitochondrial DNA contains 5-methylcytosine residues in CpG dinucleotide contexts, much like the nuclear genome, provided a simple analogy for the role this modification might play in mitochondria. Additionally, the absence of a known mitochondrial methyltransferase opened the door for one of the catalytically active nuclear-encoded enzymes to fill that role. The studies described throughout this dissertation were aimed at identifying and characterizing the DNA methyltransferase responsible for mtDNA methylation, quantifying its methylation activity, and understanding the role of DNA methylation in regulating mitochondrial function.

Chapter 2: Identification, Regulation and Function of Mitochondrial DNMT1 (mtDNMT1)

INTRODUCTION

I. Mitochondrial transcription

The D-loop regulatory region of mtDNA contains the bi-directional promoters for initiation of both heavy and light strand transcription, and serves as the site for assembly of transcription initiation complexes (Figure 2-1). Transcription of the light strand begins at a single promoter, appropriately called light-strand promoter (LSP). RNA synthesis covers the entire length of the light strand, generating a single transcript that is processed to produce 1 mRNA and 8 of the 22 tRNAs (1). Unlike the light strand, transcription of the heavy strand is initiated at two adjacent promoters, HSP1 and HSP2. Transcription starting at HSP2 generates a long, polygenic message that is processed to produce 12 mRNAs, 14 tRNAs and both of the 2 rRNAs (1). Transcription from HSP1 is dedicated solely to the production of the 2 rRNAs, to support the mitochondrion's need for high-level production of rRNA, relative to mRNA (1).

The driving force behind mitochondrial transcription is the mitochondrial RNA polymerase (POLRMT). Perhaps surprisingly, POLRMT is encoded on the nuclear genome, and gains access to the mitochondrial compartment via a cleavable, N-terminal mitochondrial targeting sequence (MTS). Related to the T7 family of bacteriophage polymerases, human POLRMT complexes with mitochondrial transcription factor B2 (mtTFB2) to form the core, two-component mitochondrial transcription machinery (Figure 2-1) (13). Other factors, including the human mitochondrial transcription factor A (TFAM), were once thought to be included in the

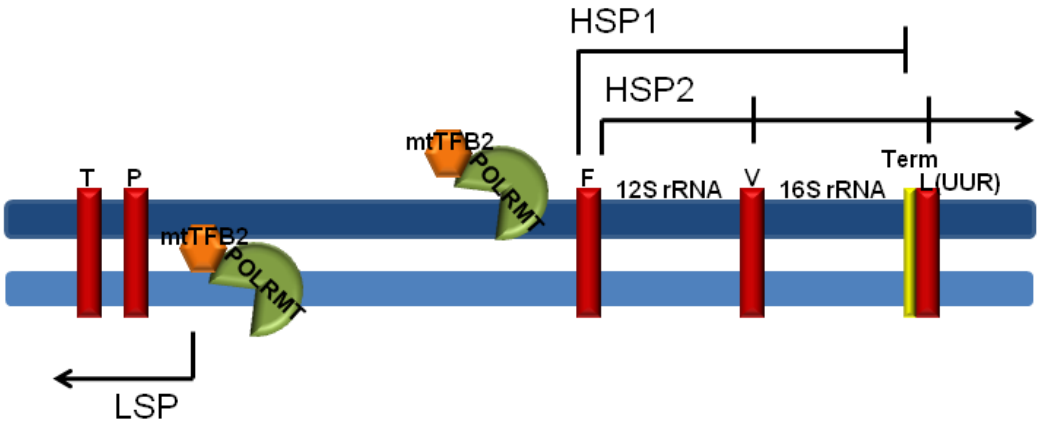


Figure 2-1: Schematic representation of mitochondrial transcription from HSP1, HSP2 and LSP. The H-strand of mtDNA is shown in dark blue; the L-strand is in light blue; tRNAs are represented by red bars and labeled by their single-letter code. Transcription initiation complexes, consisting of mitochondrial RNA polymerase (POLRMT) and mitochondrial transcription factor B2 (mtTFB2), assemble at bidirectional promoters within the D-loop. Transcripts initiating at LSP or HSP2 can traverse the entire mtDNA template, whereas those initiating at HSP1 terminate at a transcription terminator (Term, yellow bar), located within the $tRNA^{L(UUR)}$ gene. Adapted from figure by Scarpulla R (2008) *Physiol Rev* 88: 611-638.

repertoire of proteins and transcription factors comprising the transcriptional apparatus. However, it has recently been shown that TFAM is, instead, dispensable as part of the core machinery (13).

TFAM is a high-mobility-group (HMG)-box protein, also encoded on the nuclear genome, and shares homology with a group of HMG-box proteins involved in the maintenance and architecture of DNA in nuclear chromatin (84). Like HMG-box proteins in the nucleus, TFAM binds to specific recognition sites within the core promoter of mtDNA, a region which lies between the LSP and two HSPs, and functions to unwind mtDNA, making it more accessible to transcriptional machinery (1, 3). In addition to binding at specific sites, TFAM has also been reported to coat the entire mitochondrial genome, contributing to its maintenance and stabilization (1, 3, 14). TFAM knockout mice exhibit embryonic lethality paired with a drastic reduction in mtDNA content, substantiating its essential role in mtDNA maintenance (14).

Termination of transcription occurs at a strong bi-directional terminator located within the $tRNA^{Leu}$ gene, which lies between the end of the 16S rRNA gene and the first protein-coding unit on the heavy strand, NADH dehydrogenase subunit 1 (ND1) (1). This 28 nucleotide consensus site is bound by the mitochondrial termination factor 1 (mTERF1) (1). Interestingly, mTERF1 not only binds the termination site, but also HSP1, stimulating transcription, and thus coupling initiation and termination events (Figure 2-1) (1, 3). When mTERF1 binds both sites simultaneously, it results in looping-out of the DNA, forming a transcription bubble to maintain a high rate of rRNA synthesis (1).

Transcription of the mitochondrial genome is thought to be co-regulated with nuclear components of the respiratory chain complexes (1, 3). In mammals, oxidative stress results in stabilization of peroxisome proliferator activated receptor γ-coactivator 1α (PGC1α), which

activates the transcription of several nuclear-encoded transcription factors, including nuclear respiratory factor 1 (NRF1). PGC1α and NRF1 form a complex that in turn up-regulates transcription of TFAM, mtTFB1 and mtTFB2, as well as multiple members of the mitochondrial respiratory chain complexes (1, 3). This coordinated series of activation events ultimately serves to stimulate mitochondrial biogenesis as a protective cellular response (1).

II. Import of nuclear-encoded proteins into mitochondria

Of the approximately 1500 proteins present in mitochondria only about 1% originate from within the mitochondrial compartment (85). The other 99% of "mitochondrial" proteins are encoded by the nuclear genome, synthesized on ribosomes in the cytosol, and transported into mitochondria using complex translocase machinery (85). There are several major protein import pathways, each with distinct features based on the protein's ultimate submitochondrial destination, but for simplicity, our discussion will be limited to the presequence pathway (85).

The presequence pathway is the most common route of import. It relies on a specific mitochondrial targeting signal (MTS), encoded within the sequence of the transported protein, to be recognized by receptors on the outer mitochondrial membrane (Figure 2-2). The protein gains entry into the mitochondrial compartment by first passing through the translocase of the outer membrane (TOM) complex. The TOM machinery contains a highly conserved core component: the Tom40 integral membrane protein, which forms a β-barrel structure and serves as the channel through the outer membrane. The Tom20 subunit of the TOM complex recognizes the MTS of the preprotein and directs the protein to the Tom22 receptor, which then sends it through the Tom40 channel. Other, smaller Tom proteins are involved in the overall stability of the complex, and participate in the transfer of preproteins from one subunit to another (85).

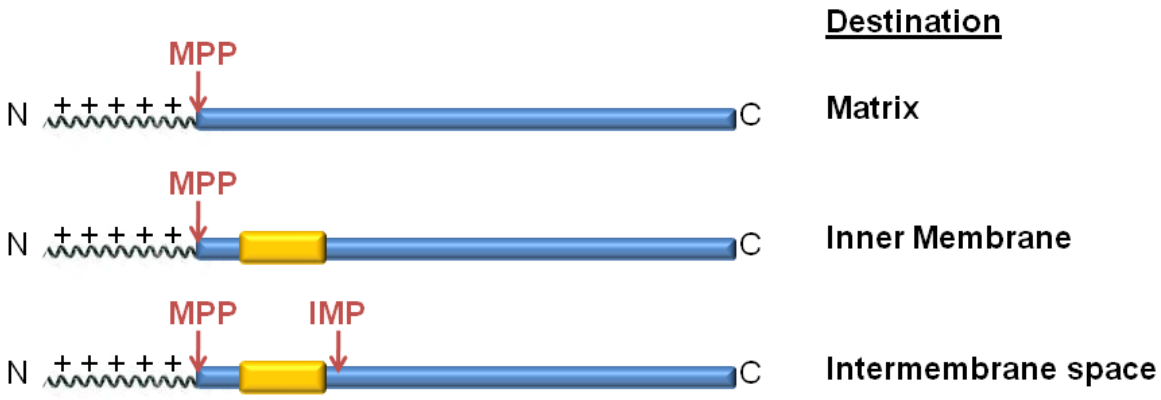


Figure 2-2: Characteristics of a mitochondrial targeting sequence (MTS). N-terminal, cleavable presequences are the classical types of mitochondrial targeting sequences, and can be grouped into three general classifications based on where they direct the attached protein within mitochondria: matrix, inner membrane, intermembrane space. MPP, mitochondrial processing peptidase; IMP, inner membrane peptidase. Adapted from figure by Chacinska A, *et al.* (2009) *Cell* 138: 628-644.

Once inside the outer membrane, the MTS becomes important for relaying information about where the preprotein should be directed. There are three general classes of mitochondrial presequences, based on the final destination of the protein: matrix, inner membrane and intermembrane space (Figure 2-2) (85). These signals are typically 15-50 amino acids in length, and form positively-charged amphipathic α helices (6, 7, 85). Upon import, most presequences are proteolytically cleaved by the mitochondrial processing peptidase (MPP). Most N-terminal presequences direct their attached proteins to the matrix, unless they contain a hydrophobic sorting signal located immediately downstream of the MTS (85). This hydrophobic sorting signal serves to arrest translocation by anchoring the precursor protein within the inner membrane. For intermembrane space proteins, the inner membrane peptidase (IMP) generally cleaves the precursor protein just downstream of the hydrophobic anchor sequence, forming a smaller mature protein (85).

Traversing the inner membrane is considerably more complicated, as it involves the translocase of the inner membrane (TIM) complex, the presequence associated motor (PAM) and two energy sources: ATP and the electrochemical membrane potential (85). Like the TOM complex, the TIM complex contains a highly conserved core module, consisting of a Tim50 receptor protein, a Tim23 channel-forming protein, and Tim17, which recruits PAM and plays a role in sorting of preproteins. Although not part of the core machinery, Tim21 also plays a critical role in import, working with Tim50 and Tim23 to aid in transient interaction between the TOM and TIM complexes, and facilitate translocation of preproteins through the outer and inner membranes (85). Once through the inner membrane, preproteins are bound by Tim44 and the central subunit of the PAM motor, mtHsp70, to be driven into the mitochondrial matrix (85). PAM requires hydrolysis of ATP to drive transfer of preproteins into the matrix. The negative

charge on the matrix side of the membrane potential also helps preproteins cross through the inner membrane by exerting an electrophoretic effect on the presequences, which carry an overall positive charge (85).

Either during or just after translocation of the preprotein into the matrix, the mitochondrial processing peptidase (MPP) removes the presequence. This allows the mature protein to fold into its active, folded form (85).

III. Regulation of DNMT1 by p53 in the nucleus

p53 is a tumor suppressor protein, encoded by one of the most frequently mutated genes in cancer. p53 functions as a DNA sequence-specific transcription factor and is highly sensitive to DNA damage and general conditions of cellular stress. When activated by stress, p53 becomes stabilized and can trigger apoptosis or p53-dependent cell cycle arrest (86). Most p53-mediated effects have been attributed to its function in the nucleus; however, p53 has also been detected in mitochondria, and proposed to influence mtDNA copy number, mitochondrial respiration and overall mitochondrial function (86-88).

Studies performed by Erica Peterson in our laboratory described a link between p53 status and regulation of DNMT1 expression (89). A sequence matching the consensus p53 recognition site was found within the genomic sequence for DNMT1, upstream of the translational start site (Table 2-1). Binding of p53 to the DNMT1 locus was demonstrated both *in vitro*, by electrophoretic mobility shift assay (EMSA), and *in vivo* by chromatin immunoprecipitation (ChIP) assays (89). Deletion of p53 in both human tumor and normal mouse cell lines induced a 3- to 6-fold increase in DNMT1 expression at both the mRNA and protein levels (89), suggesting that p53 serves to repress *DNMT1* transcription through a

Table 2-1: Consensus sequences for putative p53 binding sites within the MTS for DNMT1.

p53 consensus	RRRCWWGYYY	N(0-13)	RRRCWWGYYY	Base Pair Position
Mouse	GCGCATGCGC	15	TAGCATGGTC	-163 to -130
Human	GCGCATGCGT	7	GGGCATGGCC	-207 to -181
Chimp	GCGCATGCGT	7	GGGCATGGCC	-207 to -181
Rat	GCGCATGCGC	9	GGGCATAGCAT	-163 to -134
Cow	GCGCATGCGC	12	GAGCCAGTTCC	-205 to -173

mechanism that involves its direct DNA binding to the *DNMT1* promoter. Studies described below extended these original observations to examine the effects of p53-mediated regulation of DNMT1 in the mitochondrial compartment.

IV. Objectives

Several nuclear-encoded genes involved in mitochondrial function, including *PGC1α*, are regulated by DNA methylation (90). Conversely, it has been suggested that mitochondria are able to influence cytosine methylation levels in the nucleus by modulating the flux of one-carbon units for the generation of *S*-adenosyl-L-methionine, the methyl group donor in the DNA methylation reaction (91). Thus, epigenetic regulation of nuclear gene expression appears to have a mitochondrial component. The presence of cytosine methylation in mtDNA, without a known mitochondrial methyltransferase, led us to ask whether this epigenetic modification might play a role in the coordinated regulation of mitochondrial gene expression from both nuclear and mitochondrial genomes.

MATERIALS AND METHODS

Materials

HCT116 p53+/+ and HCT116 p53-/- cells were obtained from Dr. Bert Vogelstein, Johns Hopkins University (Baltimore, MD). Primary MEFs were prepared from E12.5-E13.5 mouse embryos. HCT116 neo^S cells were obtained from American Type Culture Collection (ATCC, Manassas, VA). Protein G sepharose beads (cat# 17-0618-01) were purchased from GE Healthcare (Uppsala, Sweden). DNA substrates used in β-glucosyltransferase assay were synthesized in the presence of dCTP, 5m-dCTP or 5hm-dCTP and purchased from Active Motif

(Carlsbad, CA). DNA oligo primers (25nmoles) were designed using OligoAnalyzer software from Integrated DNA Technologies (IDT, Coralville, IA) and purchased from Invitrogen (Carlsbad, CA), unless described otherwise. Quantitect SYBR Green Mastermix (cat# 204145) was purchased from Qiagen and used for all qPCR analysis. A Bioruptor water bath sonicator (Diagenode, Denville, NJ) was used to shear gDNA to a range of 400-700bp fragments. BioRad DNA Engine Peltier thermal cyclers, fitted with Chromo 4 Real-Time Fluorescence Detector attachments were used for all studies involving qPCR. Real-time qPCR data was analyzed with Opticon Monitor 3 Software. Beckman centrifuges (Palo Alto, CA) were used for all differential centrifugation studies: Beckman J6-MI, fitted with JS 4.2 rotor, Beckman J2-HC, fitted with JA-17 rotor, and Beckman Optima L-80 XP Ultracentrifuge, fitted with 50-TI rotor. Antibodies were purchased from Active Motif (Carlsbad, CA), Abcam (Cambridge, MA), Sigma (St. Louis, MO), Upstate Biotechnology (Billerica, MA), Santa Cruz (Santa Cruz, CA), Imgenex (San Diego, CA), Pierce (Rockford, IL), Invitrogen (Carlsbad, CA), and received as a gift from Dr. Andrew Larner (VCU), and will be discussed individually below.

Methods

Cell Culture

WT and p53-/- MEFs were grown at 37°C in 10% CO_2, in DMEM medium (Gibco/Invitrogen) supplemented with 10% FBS. HCT116neo^S cells were grown at 37°C in 5% CO_2, in RPMI 1640 medium (Gibco/Invitrogen) supplemented with 10% FBS. HCT116 p53+/+ and HCT116 p53-/- cells were grown at 37°C in 5% CO_2, in RPMI 1640 medium (Gibco/Invitrogen) supplemented with 10% FBS. All cell media were obtained from Gibco/Invitrogen and FBS was obtained from Gemini.

For experimental consistency and because DNMT1 is cell-cycle regulated (78, 92), all experiments were set up using cells treated identically in the days leading up to an experiment. Cells were fed 24 hrs prior to plating and split and seeded into appropriate dishes or flasks 24 hrs prior to harvest. Cells were harvested approximately 24 hrs after plating to ensure that the majority of cells were still in S-phase, and expressing the maximal level of DNMT1.

Subcellular fractionation – Preparation of Crude and Percoll-purified mitochondria, and cytosolic isolation

Cells from 3-4 sub-confluent 150mm dishes were washed twice with 2mL of cold 1x PBS, pH 7.4, and placed on ice. While still on ice, cells were scraped in 2mL of 1x PBS using sterile cell scrapers. Cells were transferred to a pre-weighed 15mL conical tube (BD Falcon, cat # 352096), pelleted and weighed to obtain the whole cell pellet weight. The whole cell pellet from approximately 50-60 million cells was resuspended in 3mL of mitochondrial homogenization buffer (0.25M sucrose, 10mM Tris-HCl, pH7.0, 1mM EDTA, pH 6.8) containing complete protease inhibitors (Complete EDTA-free Protease Inhibitor Cocktail tablets, Roche, cat # 11873580001, 1 tablet/25mL homogenization buffer). An aliquot (5-10%) of this resuspended cell volume was retained in a separate tube to generate the whole cell lysate, prepared by pelleting the cells (900xg, 5 minutes, 4°C), obtaining a pellet weight, and resuspending the whole cell pellet in SDS Lysis buffer (62.5 mM Tris pH 6.8, 5% glycerol, 2% SDS, 5% β-mercaptoethanol, and 1x complete protease inhibitor cocktail (Roche)) at a volume 7.5x the pellet weight. The remainder of the resuspended cell volume (90-95%) was transferred to a pre-chilled 7mL glass dounce homogenizer (Wheaton) and incubated on ice for 5 minutes to allow the cells to swell in the hypotonic homogenization buffer. Resuspended cells were

subjected to three rounds of 15 dounce strokes, followed by centrifugation at 900xg (2,200rpm) in a Beckman J6-MI centrifuge for 5 minutes at 4°C, each time separating the post-nuclear supernatant (containing mitochondrial and cytosolic fractions) from the nuclear/unbroken cell fraction. This pellet was resuspended in another 3mL of mitochondrial homogenization buffer to repeat douncing. After three rounds of douncing, the collected post-nuclear supernatant fraction was again centrifuged at 900xg for 5 minutes at 4°C to pellet any contaminating nuclear/unbroken cells. The entire post-nuclear supernatant fraction was then transferred to a 14mL round-bottom tube (BD Falcon, cat # 352059) and centrifuged at 10,000xg (9,000rpm) in a Beckman J2-HC centrifuge for 15 minutes at 4°C to pellet the mitochondria. The mitochondrial pellet was washed 2-3 times with 1mL of mitochondrial homogenization buffer, centrifuging at 10,000xg in a tabletop microfuge for 5 minutes at 4°C. The final mitochondrial pellet was weighed, and will be referred to as a "Crude" mitochondrial preparation.

A Percoll gradient was constructed; 6.55mL Percoll (Amersham, cat # 17-0891-02) was mixed with 2.2mL of 2.55M sucrose, 12.25mL of TE, pH 7.4, and ddH2O to a total volume of 21mL, in a 50mL conical tube. This solution was mixed well, and 10mL was aliquoted into two 50TI Beckman Ultracentrifuge tubes (cat # 344085). The crude mitochondrial pellet was resuspended in 1mL of mitochondrial homogenization buffer with complete protease inhibitors, and overlaid onto the Percoll by gently pipetting the resuspended mitochondria down the side of the tube. Samples were centrifuged at 60,000xg (30,000rpm) in a Beckman Ultracentrifuge for 45 minutes at 4°C. Following high-speed centrifugation, mitochondria could be seen to form a hazy white band that migrated near the center of the gradient. This band was conservatively transferred into new 14mL round-bottom tubes, and washed 2-3 times with 2-3mL of mitochondrial homogenization buffer, spinning at 10,000xg for 15 minutes at 4°C in the

Beckman J2-HC after each wash. The final Percoll-purified mitochondrial pellet was weighed, and will be referred to as a "Percoll-purified" mitochondrial preparation.

The cytosolic fraction was obtained by collecting the post-mitochondrial supernatant following high-speed centrifugation to pellet mitochondria (10,000xg, 15 minutes, 4°C in Beckman J2-HC, described above). An equal volume of ice-cold 20% TCA was added to the post-mitochondrial supernatant, mixed well, and incubated on ice for 30 minutes. Cytosolic proteins were pelleted at 6,000xg (7,500rpm in Beckman J2-HC centrifuge) for 15 minutes at 4°C, washed 2-3 times with 3-5mL of ice-cold acetone, each time spinning at 7,500rpm for 5 minutes at 4°C to collect the pellet. After the last wash, the cytosolic pellet was allowed to air-dry on ice for approximately 5 minutes, before being dissolved in SDS lysis buffer (described below) for direct analysis by immunoblot.

Methyl-DNA Immunoprecipitation (MeDIP) using purified mtDNA

The specificity of both 5mC and 5hmC antibodies for their respective cytosine modifications was verified by Prashant Thakkar using defined DNA substrates synthesized in the presence of dCTP, 5m-dCTP, or 5hm-dCTP (Active Motif).

Mitochondria were isolated from 50-60 million HCT116 cells, as described above. The mitochondrial pellet was washed twice with 1mL of mitochondrial homogenization buffer with complete protease inhibitors (Complete EDTA-free Protease Inhibitor Cocktail tablets, Roche), and once with 1mL of mitochondrial homogenization buffer in the absence of protease inhibitors. To washed mitochondrial pellets, formaldehyde crosslinking solution (11% formaldehyde, 50mM HEPES, pH 8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA) was added to 1/10th the final volume, and incubated for 15min at 37°C. Crosslinking was terminated by adding glycine

to a final concentration of 125mM and incubating at room temperature for 10min. Crosslinked mitochondria were pelleted by centrifugation at 10,000xg for 5min at 4°C, and washed once with 1mL of cold 1x PBS, pH 7.4. Mitochondrial pellets were split into two 1.5mL microcentrifuge tubes and re-pelleted by centrifugation, then lysed in 1mL mtIP Lysis buffer (0.5% Triton X-100, 300mM NaCl, 50mM Tris-HCl, pH 7.4, 100µg/mL leupeptin, 200µM PMSF, complete protease inhibitor cocktail). The mitochondrial lysate was pipeted up and down several times to mix thoroughly and incubated on ice for 15min with occasional inversion of the tubes, to allow complete lysis.

Crosslinked mitochondrial DNA was sonicated using the Diagenode Bioruptor water bath sonicator at the highest pulse, cycling between 30sec on and 30sec off, for two rounds of 15min each. DNA fragmentation was assessed on a 1% TAE-EtBr agarose gel to verify proper shearing to an average of 400bp. Sonicated DNA (10µL) was run alongside 1µg of unsonicated gDNA. Crosslinks were reversed by adding 0.5% SDS and incubating samples overnight at 65°C. Sonicated DNA was precipitated by adding $1/10^{th}$ volume (30µL) of 3M NaOAc, 30µg of UltraPure Glycogen (Invitrogen, cat # 10814010), and 2 volumes (600µL) of 100% ethanol per sample, followed by incubation at -80°C for approximately 2 hrs. Precipitated DNA was centrifuged at high speed in a microfuge for 15 minutes at 4°C. The DNA pellet was washed once with 1mL of 70% ethanol and allowed to air dry on the benchtop for approximately 5 minutes. DNA pellets were resuspended in 200µL HPLC H2O with 50µg/mL of RNase A (Sigma, cat # R5503), and samples were incubated at 37°C for 1 hr. 100µg/mL Proteinase K (1µL of 20mg/mL stock, Bioline, cat # BIO-37037) and 0.25% SDS were added to each tube, and mixed well. DNA samples were incubated at 37°C overnight for complete digestion of RNA and protein contaminants. Digested DNA was subjected to two rounds of

phenol/chloroform/isoamyl alcohol extraction (Invitrogen, cat # 15593031), and precipitated as described above. Precipitated DNA was resuspended in 500µL of 10mM Tris, pH 8.0. DNA was quantitated using the ND-1000 NanoDrop spectrophotometer, and 4µg purified mtDNA was used as input into each MeDIP assay. The DNA was denatured by boiling for 10 minutes, and then immediately cooled on ice for 10 minutes. An aliquot (51µL) of 10x IP buffer (100mM NaPO4, pH 7.0, 1.4M NaCl, 0.5% Triton X-100) and 2µg of each antibody for IP were added to the denatured DNA sample, and incubated at 4°C overnight with end-over-end rotation. A total of three conditions were examined for each DNA sample: Nonspecific IgG (-) control (Millipore, cat # 12-371), 5mC antibody (Active Motif, cat # 39649), and 5hmC antibody (Active Motif, cat # 39770).

Lambda DNA (New England BioLabs, cat # N3011L) was sonicated using a Diagenode Bioruptor water bath sonicator at the highest pulse for 2 cycles of 15 minutes each, with 1 minute on and 1 minute off. After each cycle, the tubes were removed from the sonicator, tapped to mix and quickspun to collect all DNA at the bottom of the tubes.

Protein G sepharose beads (GE Healthcare, cat # 17-0618-01) at a volume of 20µL per sample, per condition, were pipeted using a 2-200µL tip with the end cut off and pelleted by centrifugation at 4,500rpm for 5 minutes at 4°C in a tabletop microfuge, to remove the 70% ethanol resuspension mix. The beads were washed with 400µL of 1x PBS-0.1% BSA solution three times, incubating the beads with rotation for 5 minutes at 4°C and spinning at 4,500rpm for 5 minutes at 4°C. At the end of the third wash, the volume of pelleted beads was estimated, and an equal volume of 1x IP buffer was added to make a 50% bead slurry. BSA (5µg per 30µL of 50% slurry) and sonicated lambda DNA (5µg per 30µL of 50% slurry) were added to the bead slurry and incubated for 3 hours at 4°C with rotation to block the beads and prevent non-specific

IgG binding. After blocking, the beads were again washed with 1x PBS-0.1% BSA as described above, and the pellet was resuspended in an equal volume of 1x IP buffer, and mixed by tapping the tube.

Approximately 20µL of blocked, washed bead slurry was added to each tube containing a DNA-antibody complex. These samples were incubated for 2 hours at 4°C with rotation to allow DNA-antibody-Protein G bead complexes to form. Beads were collected by centrifuging at 4,500rpm for 5 minutes at 4°C, and washed three times, each with 500µL 1x IP buffer for 10 minutes at 4°C with rotation, and spinning at 4,500rpm for 2 minutes at 4°C. After the third wash, the beads were resuspended in 250µL of Proteinase K digestion buffer (50mM Tris, pH 8.0, 10mM EDTA, 0.5% SDS), and mixed by tapping the tube. Seventy µg of Proteinase K (3.5µL of 20mg/mL stock, Bioline, cat # BIO-37037) was added to each sample, and incubated for 3 hours at 50°C with shaking to disrupt the antibody-Protein G bead interactions. Immunoprecipitated DNA was extracted with 250µL of Phenol:Chloroform:Isoamyl alcohol (25:24:1), and precipitated using $1/10^{th}$ volume (25µL) of 3M NaOAc, 40µg of UltraPure Glycogen (2µL of 20µg/µL stock, Invitrogen, cat # 10814010), and 3 volumes (750µL) of 100% ethanol per sample. The DNA pellet was allowed to air-dry at room temperature for approximately 5 minutes, resuspended in 75µL TE buffer, pH 8.0, and stored at -20°C.

Quantitative PCR using Quantitect SYBR Green mastermix and mitochondrial-specific primers (see Table 2-2 for primer sequences and Figure 2-2 for a map of primer positions along mtDNA) was performed to measure the level of enrichment of mtDNA sequences in immunoprecipitated material pulled down by each antibody: anti-5mC, anti-5hmC or nonspecific IgG. Absolute amounts of immunoprecipitated mtDNA were quantitated using a standard curve of purified mtDNA, and expressed relative to input.

Table 2-2: Mitochondrial primer sequences.

Species	Primer Set	Forward	Reverse
		Amplification of Mitochondrial Targeting Sequences	
Mouse	mMTS	5'-ATGCGCACTCCCTTCGGGCATAG-3'	5'-CTTGCAGGTTGCAGACGACAG-3'
Human	hMTS	5'-ATGAATGAATGCCTCGGGCAC-3'	5'-CTCGGAGGCTTCAGCAGACGC-3'
		Quantitation of DNMT1 Abundance	
Mouse	mtDNMT1	5'-ACTCTCTTGCCCTGTGTGGTACATG-3'	5'-TCTTTCCAAGTCTTTGAGCCGCC-3'
	Total DNMT1	5'-TCAGAGCTGTTCTGTCGTCTGCAA-3'	5'-TCTTTCCAAGTCTTTGAGCCGCC-3'
	β-Actin	5'-GACCCAGATCATGTTTGAGACC-3'	5'-ATCAGAATGCCTGTGGTACGAC-3'
Human	mtDNMT1	5'-TCCCTGGGCATGGCCGGCTC-3'	5'-CTCTTTCCAAATCTTGAGCCGCC-3'
	Total DNMT1	5'-TCCGAGATGCCGGCGCGTACC-3'	5'-CTCTTTCCAAATCTTGAGCCGCC-3'
	β-Actin	5'-CCACGAAACTACCTTCAACTCC-3'	5'-TCATACTCCTGCTTGCTGATCC-3'
		Mitochondrial Gene Expression	
Mouse	ND1	5'-CAGGATGAGCCTCAAACTCCA-3'	5'-CGGCTCGTAAAGCTCCGA-3'
	ATP6	5'-ATTCCCATCCTCAAAACGCC-3'	5'-TGTTGGAAAGAATGGAGACGGT-3'
	ND6	5'-AGTTATGTTGGAAGGAGGGATTGG-3'	5'-TACCCGCAAACAAAGATCACCCAG-3'
	COX1	5'-TCGCAATTCCTACCGGTGTC-3'	5'-CGTGTAGGGTTGCAAGTCAGC-3'
	18S rRNA	5'-GTCTGTGATGCCCTTAGATG-3'	5'-AGCTTATGACCCGCACTTAC-3'
		Mitochondrial MeDIP (5mC and 5hmC)	
Human	ATP6	5'-ATTCAACCAATAGCCCTGGCCG-3'	5'-ACGTAGGCTTGGATTAAGGCGAC-3'
	COX2	5'-ACAGATGCAATTCCCGGACGTC-3'	5'-TGGGCATGAAACTGTGGTTTGCTC-3'
	12S	5'-AGTTCACCCTCTAAATCACCACG-3'	5'-TGACTTGGGTTAATCGTGTGACC-3'
	16S-1	5'-ACCTTACTACCAGACAACCTTAGCC-3'	5'-TAGCTGTTCTTAGGTAGCTCGTCTGG-3'
		qPCR Quantitation of 5hmC	
Mouse	16S-2	5'-AAAAGAGGGACAGCTCTTCTGGAACG-3'	5'-TCGTTTAGCCGTTCATGCTAGTCC-3'
	16S-3	5'-CGGTTTCTATCTATTTACGATTTCTCCC-3'	5'-GCCACCCTAATAACCTTCTCTAGG-3'
	ATP6/COX3	5'-CCCACCAACAGCTACCATTAC-3'	5'-CTAGACCTGATGTTAGAAGGAGGG-3'
Human	ATP6	5'-ATTCAACCAATAGCCCTGGCCG-3'	5'-ACGTAGGCTTGGATTAAGGCGAC-3'
	12S	5'-AGTTCACCCTCTAAATCACCACG-3'	5'-TGACTTGGGTTAATCGTGTGACC-3'
	2	5'-GAGACAAGTCGTAACATGGTAAG-3'	5'-GGGTAAGGTTTGCCGAGTTCCT-3'

β-Glucosyltransferase Assay – Relative quantitation of 5hmC levels

The sequence-specific detection of 5hmC at *Gla*I restriction sites was determined using the Quest 5hmC Detection kit (Zymo Research, cat # D5416), according to manufacturer's specifications. *Gla*I restriction enzyme was purchased separately from SibEnzyme (Russia, cat # E494), supplied with 10x SEB Reaction Buffer. Either purified mtDNA or total cellular DNA (80ng) was used as starting template. Unmethylated, fully methylated and fully hydroxymethylated control DNAs supplied by Active Motif were used to validate the assay. The assay relies on the fidelity of a series of modifying and restriction enzymes, and the sensitivity of qPCR to detect even small differences in the availability of full-length template DNA.

Template DNA from HCT116 cells or MEFs was first incubated with the bacterial methylase, *M.Sss*I (New England Biolabs, cat # M0226S) to methylate all unmethylated CpG dinucleotides, and GpC methyltransferase *M.CviP*I (New England BioLabs, cat # M0227S) to methylate all GpC dinucleotides within the template. Then, DNA was incubated with β-glucosyltransferase (β-gt) to add a bulky glucosyl moiety to all naturally-occurring 5hmC residues. Finally, template DNA was digested with *Gla*I enzyme, which cleaves only when the cytosines within its restriction site are either fully methylated or fully hydroxymethylated. *Gla*I cannot cleave its target sequence when it has been glucosylated, so modification by β-gt affords the protection of a 5hmC residue from *Gla*I cleavage. Modified, digested DNA was used as template in qPCR with primers specific to various regions along the mitochondrial genome (see Table 2-2). The level of enrichment of 5hmC was normalized to the input signal for each amplicon.

Bioinformatic analyses of the DNMT1 mitochondrial targeting sequence

The mitochondrial targeting sequences for DNMT1 were located using UCSC Genome Browser software (http://genome.ucsc.edu/) by querying the reference sequence ("RefSeq") of DNMT1 for the promoter/upstream region as far as 500 bases upstream of the coding sequence. In all species for which the 5' flanking sequence of DNMT1 was available, the retrieved genomic sequence was copied and pasted into the ExPasy Translate tool (http://www.expasy.ch/tools/dna.html) to convert the nucleotide sequence into its predicted protein sequence. Once translated, the appropriate 5' to 3' frame was selected such that the sequence downstream of the nuclear translational start (labeled ATG3 in Figure 2-6) matched the known protein sequence for DNMT1 (… MPARTAPARV …). The correct frame of the entire translated sequence for DNMT1, starting at each upstream ATG, was consecutively copied and pasted into the MitoProt II query box (http://ihg.gsf.de/ihg/mitoprot.html) to predict whether a protein originating at each of the putative upstream start sites translocates to the mitochondria. Mitochondrial targeting sequences for DNMT1 in all available species were analyzed for the presence of transcription factor binding sites using MatInspector software (http://www.genomatix.de/en/index.html).

Cloning of MTS and confocal microscopy

The mitochondrial targeting sequences for both human and mouse DNMT1 were amplified by PCR, against random-primed cDNA templates, starting at ATG1 and ending just upstream of ATG3 using primers as described in Table 2-2. PCR products were separated on a 1% TAE agarose gel stained with ethidum bromide, and visualized by UV transillumination. The desired PCR products (~160bp for mouse, ~300bp for human) were excised from the gel using a

clean razor blade, the gel slice was weighed, and the PCR product was purified using the Wizard SV gel and PCR Clean-up System (Promega, cat # A9282), according to manufacturer's instructions. Purified PCR products were cloned into the Vivid Colors pcDNA 6.2/EmGFP mammalian expression vector (Invitrogen, cat # K359-20) with an in-frame C-terminal GFP tag. Dr. Shirley Taylor transfected these vectors into NIH/3T3 mouse lung fibroblasts or HCT116 human colon carcinoma cells grown on poly-$_{L}$-lysine-coated glass coverslips. Transfections were performed using Polyjet liposomes (SignaGen, cat # SL100688) for HCT116 cells, or Amaxa MEF1 Nucleofector Kit (Lonza, cat # VPD-1004) for NIH/3T3 cells. Both transfection reagents were used according to the manufacturer's instructions. GFP fluorescence was observed at 24hrs post-transfection to verify expression of the GFP-tagged constructs and to estimate transfection efficiency. At 48hrs post-transfection, Dr. Shirley Taylor fixed cells with 4% paraformaldehyde and stained with 25nM MitoTracker Red CMXRos (Molecular Probes, M-7512) for 45 minutes at 37°C. Stained coverslips were washed 3 times with 1x PBS, pH 7.4 and mounted onto glass slides with ProLong Gold antifade reagent with DAPI (Invitrogen, cat # P-36931). Mitochondrial localization was visualized using a Leica TCS-SP2 AOBS confocal scanning microscope. All confocal microscopy was performed by Dr. Shirley Taylor.

RT-PCR to detect transcript for *mtDNMT1*

Total RNA was isolated from both HCT116 cells and MEFs using TRIzol Reagent (Invitrogen, cat # 15596-026), according to manufacturer's instructions. RNA was treated with amplification-grade DNaseI (Invitrogen, cat # 18068015) to remove any gDNA contamination, and reverse-transcribed with the SuperScript III First-Strand Synthesis system (Invitrogen, cat # 18080051) using oligo-d(T) primers, following the manufacturer's protocols. Endpoint PCR to

detect a mature transcript for the mitochondrial isoform of DNMT1 was performed using HotStarTaq Plus Master Mix (Qiagen, cat # 203645). Sense primers placed across ATG1 (mouse) or ATG2 (human) were paired with an antisense primer that crossed the exon 1-2 boundary by four nucleotides (Ex1-2 asn Mod, Table 2-2). Minus RT template control reactions were run alongside +RT template reactions for each primer set. PCR products were separated on a 1% TAE agarose gel stained with ethidium bromide and visualized by UV transillumination.

SDS-PAGE and immunoblotting of subcellular fractions

Each subcellular fraction (whole cell lysate, cytosolic and mitochondrial) was collected in a pre-weighed 1.5mL microcentrifuge tube, and weighed to calculate the appropriate volume of SDS Lysis Buffer (62.5mM Tris-HCl, pH 7.5, 5% glycerol, 2% SDS, 5% β-mercaptoethanol, 1x complete protease inhibitor cocktail (Roche)) to add to each sample. SDS Lysis buffer was added at 7.5x the whole cell pellet weight, and 10x the cytosolic and mitochondrial pellet weights. Whole cell and cytosolic lysates were passed through a 21G needle fitted with 1mL syringe approximately 20 times to shear the gDNA and create a homogeneous mixture. Lysates were aliquoted and stored at -80°C until use.

Protein concentrations were calculated using a Bradford Protein assay (BioRad, cat # 500-0006) against a standard curve generated by absorbance values measured for known concentrations of BSA (Fisher, cat # 23210) read at 595nm. Protein was loaded onto SDS-PAGE gels to approximate equal cell equivalents, so that an equal signal for each compartment-specific antibody was obtained (whole cell lysate: 75µg, cytosolic lysate: 25µg, mitochondrial lysate: 18µg). The appropriate amount of total protein from each cell fraction was diluted in an equal volume of Laemmli sample buffer (BioRad, cat # 161-0737) plus 5% β-mercaptoethanol, boiled

for 5 minutes, and quickly spun to collect the samples. SDS-PAGE 4-15% gradient gels (BioRad 4-15% Tris-HCl Ready gel, cat # 161-1158) were assembled in a Mini-PROTEAN electrophoretic box unit (BioRad, model # 422 ElectroEluter), and the tank was filled with 1L of 1x running buffer (25mM Tris base, 250mM glycine, 0.1% SDS). Boiled lysate samples were loaded into the wells alongside 10µL of the Precision Plus Dual Color pre-stained protein ladder (BioRad, cat # 161-0374). Gels were resolved by electrophoresis at 150V for approximately 1hr.

Resolved proteins were transferred onto an Immobilon PVDF membrane (Millipore, cat # IPVH00010), soaked in 100% methanol and sandwiched between 15 sheets of pre-cut Whatman paper (3mm Chromatography paper, cat # 3030 917). Each sheet of Whatman paper was soaked in 1x Towbin buffer (25mM Tris-HCl, pH 8.3, 192mM glycine, 10% methanol, 0.04% SDS) and rolled smooth to remove any bubbles. The transfer sandwich was assembled in the center of the bottom plate of the Hoefer Semi-dry Transfer apparatus (model # TE77). The top of the semi-dry transfer apparatus was fitted squarely on top of the transfer sandwich and plugged into the receptacle on the bottom unit. The transfer apparatus was connected to a power supply (BioRad PowerPac Basic) and run at 80mA for 4hrs to transfer all proteins onto the PVDF membrane.

To prevent any nonspecific binding of the antibodies, the membrane was blocked in T20 StartingBlock blocking buffer (Fisher, cat # 37543), either for 1hr at room temperature or overnight at 4°C, with gentle rocking on a platform shaker. Blots were washed with 1x TBS-T (0.5M Tris-HCl, pH 7.5, 0.14M NaCl, 2.7mM KCl and 0.1% Tween 20) three times for 5 minutes each with vigorous shaking. Primary and secondary antibodies were diluted in StartingBlock blocking buffer under the conditions optimized for each antibody (see Table 2-3). Membranes were incubated in primary and/or secondary antibody for 1hr at room temperature

Table 2-3: Optimal conditions for antibodies used in mitochondrial immunoblots.

Antibody	Manufacturer	Species	Blocking Buffer	Primary Dilution	Secondary Dilution	Secondary Antibody	Protein Size (kDa)
N-DNMT1	Abcam	rabbit	StartingBlock	1:1000	1:10,000	goat anti-rabbit	185
VDAC	Sigma	rabbit	StartingBlock	1:4000	1:10,000	goat anti-rabbit	32
Tubulin	*a gift from Dr. Larner	mouse	5% milk/TBS-T	1:10,000	1:10,000	goat anti-mouse	60
H3K4me3	Upstate Biotechnology	rabbit	StartingBlock	1:1000	1:10,000	goat anti-rabbit	17
DNMT3a	Imgenex	mouse	StartingBlock	1:250	1:10,000	goat anti-mouse	120
DNMT3b	Imgenex	mouse	StartingBlock	1:250	1:5000	goat anti-mouse	110
GFP	Invitrogen	rabbit	StartingBlock	1:1000	1:10,000	goat anti-rabbit	varies

with gentle shaking. Following both primary and secondary antibody incubation, membranes were washed with 1x TBS-T three times for 10 minutes each with vigorous shaking. All secondary antibodies were conjugated with horseradish peroxidase to allow for chemiluminescent detection using the SuperSignal West Pico (Pierce, cat # 34080) and West Dura (Pierce, cat # 34075) Chemiluminescent Substrate kits according to manufacturer's instructions. Blots were developed using autoradiography film (ISC Bioexpress, cat # F-9024-8X10) and a Konica SRX-101A developer.

Regulation of mtDNMT1 expression

NRF1 and PGC1α

Murine cDNA encoding NRF1 was obtained from ATCC (Manassas, VA) and re-cloned into pDEST26/C-FLAG with an in-frame C-terminal Flag-tag. The plasmid containing PGC1α was a gift from Dr. Gregorio Gil (Virginia Commonwealth University). HCT116 cells were transfected using Polyjet liposomes (SignaGen, cat # SL100688) according to manufacturer's instructions. MEFs were transfected using the Amaxa MEF1 Nucleofector kit (Lonza, cat # VPD-1004) according to manufacturer's instructions. All cells were harvested 48 hrs after transfection, and fractionated as described above, to make whole cell and mitochondrial lysates. Aliquots of whole cell lysate (75µg) and mitochondrial lysate (18µg) were separated on 4-15% gradient gels and transferred by semi-dry transfer method, exactly as previously described. Blots were cut twice horizontally: once at the 50kDa marker, and once between the 20kDa and 25kDa markers. Each strip was probed individually with the following antibodies: 50-250kDa with the Abcam N-DNMT1 antibody (185kDa full-length DNMT1), 25kDa-50kDa with the VDAC

antibody (Pierce, 32kDa), and 0-25kDa with the H3K4me3 antibody (Upstate Biotechnology, 17kDa). Blots were probed, washed, and developed as described previously.

p53

Total RNA was isolated and reverse transcribed as previously described from WT or p53 -/- MEFs. RT-qPCR was performed using Quantitect SYBR Green mastermix, and the same primer sets as used in the studies to detect mtDNMT1 transcript in mouse (Table 2-2). The ATG1 sense primer was paired with the Ex1-2 asn Mod primer to measure abundance of mRNA encoding only mitochondrial *DNMT1*, while the ATG3 sense primer was paired with the Ex1-2 asn Mod primer to measure the abundance of total *DNMT1* mRNA. Reactions were run in triplicate and values were normalized to mouse β-actin.

WT and p53-/- MEFs were fractionated to generated whole cell, cytosolic and mitochondrial lysates, as described previously. 75µg of whole cell lysate, 25µg of cytosolic lysate and 18µg of mitochondrial lysate were separated on two identical 4-15% gradient gels and transferred to PVDF membranes as previously described. The first blot was cut once horizontally at the 25kDa marker, to probe the top portion (25-250kDa) for N-DNMT1 (Abcam), and the bottom portion (<25kDa) for H3K4me3 (Upstate Biotechnology). The second identical blot was cut once horizontally between the 37kDa and 50kDa markers, to probe the top portion (50-250kDa) with a tubulin antibody (gift from Dr. Andrew Larner, VCU, 60kDa), and the bottom portion (<50kDa) with a VDAC antibody (Pierce). Blots were probed, washed and developed as previously described.

Effects of mtDNMT1 over-expression on mitochondrial transcription

Total RNA was isolated from WT and p53-/- MEFs using TRIzol Reagent, and reverse transcribed with the SuperScript III First Strand Synthesis System, both according to the manufacturer's instructions. In the first set of experiments, cDNA was synthesized using random hexamers to obtain a more complete representation of the transcribed genome. In the second set of experiments, strand-specific cDNA was synthesized using complementary primers located within the ATP6 gene: ATP6sn for light strand cDNA synthesis, and ATP6asn for heavy strand cDNA synthesis. For both experiments, changes in mitochondrial gene expression between the two cell lines were measured using primers specific to four different regions of the mitochondrial genome, 3 on the heavy strand: ATP6, COX1 and ND1, and one on the light strand: ND6. Primer sequences are listed in Table 2-2, and their location is mapped on the mitochondrial genome in Figure 2-2. An Analysis of Variance (ANOVA) using least-square means was applied to determine the statistical significance of differences in gene expression. A random-effects ANOVA was performed using the statistical package JMP 7.0 (SAS). The least-squares mean for each gene in WT and p53-/- MEFs was obtained, along with standard errors. Each dataset was normalized to 18S rRNA expression. Three independent sets of biological samples were analyzed for each gene, with triplicate technical repeats in each sample. The ANOVA model included technical replicates as nested effects and biological replicates as random effects. The corresponding 95% confidence intervals (C.I.) were obtained using statistical methods for transformations (Delta method). Standard deviations (SDs) were computed using the formula:

$$SD = AVG_{sample}/AVG_{input} \times (\sqrt{((SD_{sample}/AVG_{sample})^2+(SD_{input}/AVG_{input})^2)}).$$

RESULTS

5mC and 5hmC modifications are present in mtDNA

Early studies using thin-layer chromatography and nearest-neighbor analysis reported the presence of 5mC in mtDNA (15-17), but there was doubt as to the purity of the mitochondrial preparations and the sensitivity of the methods used in these early reports. Here, with newer and more sensitive techniques, we not only confirm the presence of 5mC modifications in mtDNA, but we also report for the first time, the existence of 5hmC-modified residues in mtDNA (Figure 2-3). MeDIP with antibodies specific to 5mC or 5hmC modifications was performed on mtDNA made from isolates of highly-purified mitochondria. The immunoprecipitated DNA was probed by qPCR to determine the presence and relative abundance of both types of modified bases within DNA sequences specific to the mitochondrial genome. DNA immunoprecipitated with the 5mC antibody was enriched 10- to 20-fold for mtDNA over nonspecific IgG controls, across all regions tested. Interestingly, we also observed a very strong (38- to 580-fold) enrichment of mtDNA in material immunoprecipitated with the 5hmC antibody, relative to nonspecific IgG controls. The apparent lower level of enrichment of mtDNA in 5hmC-immunoprecipitated samples measured by primer 27 is likely a reflection of less efficient amplification of a longer amplicon (833bp) from mtDNA sheared to an average of 400bp. However, the enrichment of mtDNA in these samples is still a substantial 38-fold over nonspecific IgG controls. These were extremely exciting results, as they suggested that earlier studies likely underestimated both the degree and significance of epigenetic modification in the mitochondrial genome.

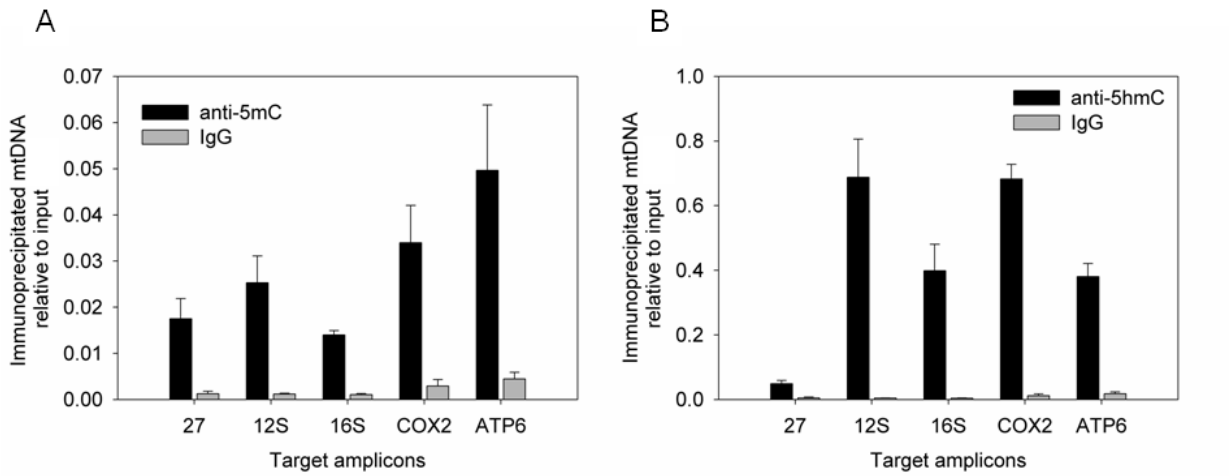


Figure 2-3: qPCR detection of 5mC and 5hmC in mtDNA. (A) mtDNA immunoprecipitation using an antibody directed against 5mC shows a 10- to 20-fold enrichment compared with IgG controls. (B) mtDNA immunoprecipitation using an antibody directed against 5hmC demonstrates substantial levels of this modified base in mtDNA (38- to 580-fold over IgG controls). (A and B) Mitochondria were purified from HCT116 human colon carcinoma cells. Mitochondrial-specific amplicons are labeled in Figure 1-2; Primers are listed in Table 2-2. Prashant Thakkar performed all MeDIP and hydroxy-MeDIP analyses.

β-Glucosyltransferase assay validates and quantifies 5hmC levels in mtDNA

We performed an independent assay to validate the presence of 5hmC residues in mtDNA (Figures 2-4 and 2-5). The Quest 5hmC Detection kit from Zymo Research provided a straightforward method for gaining quantitative information on the relative abundance of 5hmC residues in a sequence-specific manner. We used phage T4 5hmC-β-glucosyltransferase (β-gt) to determine the presence of 5hmC at *Gla*I restriction endonuclease cleavage sites. A series of control experiments using defined DNA templates containing only unmethylated cytosine, 5mC or 5hmC, confirmed that the *Gla*I restriction enzyme cleaved only sites that were fully methylated or fully hydroxymethylated, but not those containing a glucosylated 5hmC moiety (Figure 2-4). Thus, protection of mtDNA sequences from *Gla*I cleavage was afforded by glucosylation of 5hmC residues by the β-gt enzyme, and this protection could be measured by our ability to amplify across the *Gla*I cleavage site by PCR. The level of protection from *Gla*I cleavage was assessed first by endpoint PCR, and then by qPCR (Figure 2-5). 5hmC modifications were detected in three different amplicons from human mtDNA and two amplicons from both mouse mtDNA and mouse gDNA. Amplicons containing two *Gla*I restriction sites each (amplicons ATP6, 12S and 16S-3) showed approximately 50% protection in comparison to amplicons with only a single *Gla*I cleavage site (amplicons 2 and 16S-2). This suggests that there is a similar level of 5hmC at all restriction sites tested, since sequences containing two cleavage sites could only be amplified with half the frequency as those containing just one *Gla*I site. As an internal control, a mouse amplicon devoid of *Gla*I sites (ATP6/COX3) did not show protection from *Gla*I cleavage, even after 5hmC glucosylation.

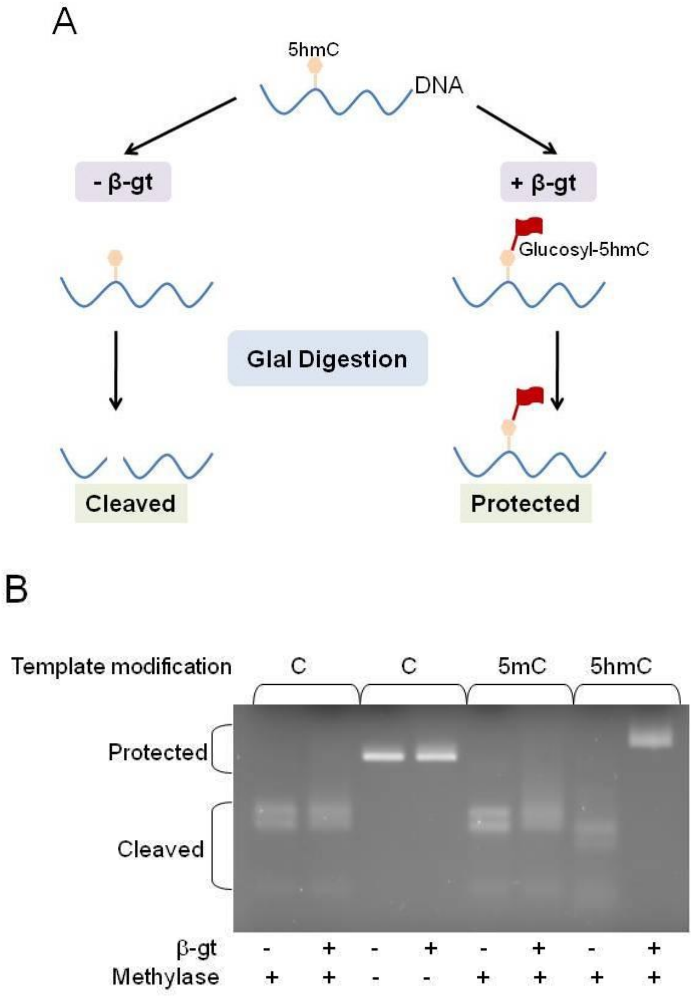


Figure 2-4: β-glucosyltransferase assay for site-specific detection of 5hmC in mtDNA. (A) Flowchart showing steps of β-glucosyltransferase assay. (B) Control DNA samples from the Adenomatous polyposis coli (APC) promoter were amplified in PCR reactions containing dCTP, 5m-dCTP or 5hm-dCTP and used to validate the specificity of the Quest 5hmC Detection kit (Zymo Research). DNA (150ng) was incubated in the presence or absence of β-glucosyltransferase (β-gt) and methylase mixture (*M.Sss*I and *M.CviP*I) as indicated, followed by *Gla*I cleavage. Products of the reaction were resolved on a 3% NuSieve agarose gel. *Gla*I cleavage was specific for methylated or hydroxymethylated DNA and was prevented by glucosylation of 5hmC residues (lanes 7 and 8). Complete methylation by the methylase mixture resulted in complete cleavage (lanes 1 and 2).

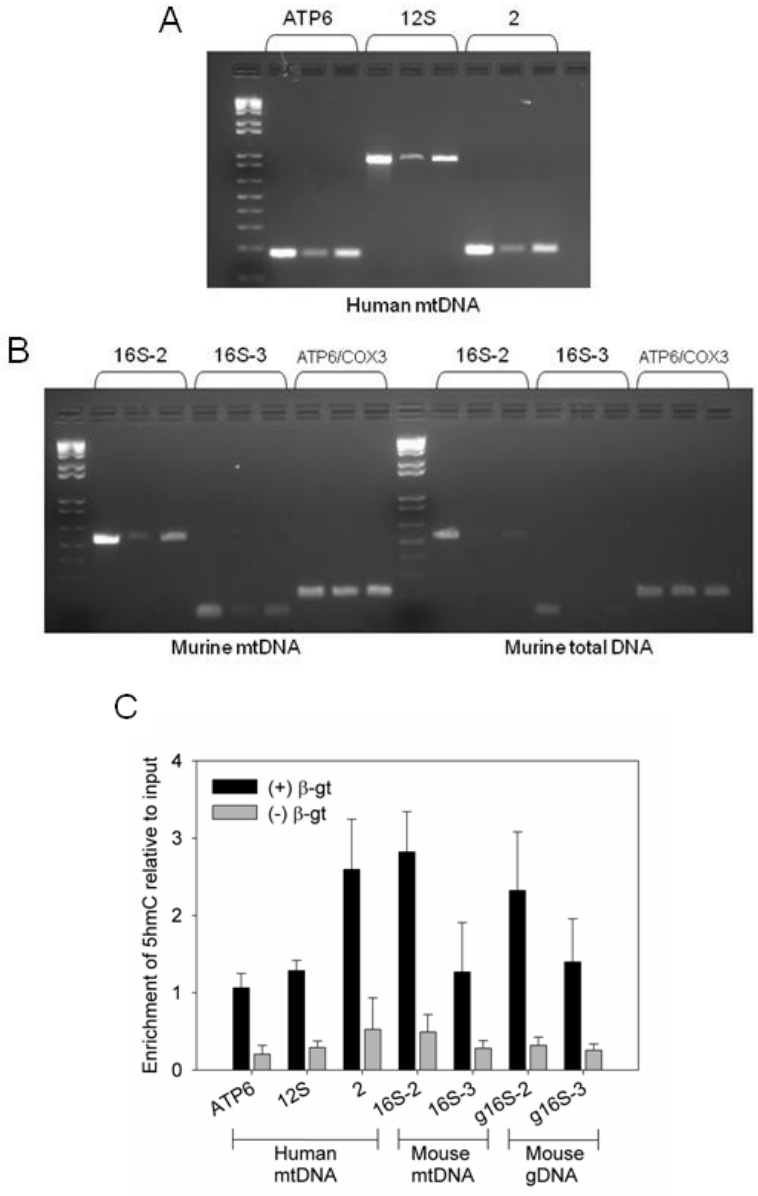


Figure 2-5: Quantitation of site-specific 5hmC levels in mtDNA. (A and B) Endpoint PCR demonstrates protection from *Gla*I cleavage of either 80ng of human mtDNA (A), or 80ng of mouse mtDNA or mouse total gDNA (B), following incubation with β-gt and methylase mixture. For each primer set, lane 1 represents input DNA, lane 2 represents (-) β-gt control and lane 3 represents (+) β-gt protection. Mouse amplicon ATP6/COX3 was expected to contain a single *Gla*I restriction site, but showed protection both in the absence and presence of β-gt. DNA sequence analysis verified the presence of a SNP deleting the restriction site (ACGT to ATGT). The data demonstrate that this assay is applicable to both purified mtDNA and total genomic DNA. (C) Enrichment of 5hmC in DNA samples was measured in triplicate and the average enrichment was graphed relative to the input for each amplicon. Error bars signify the standard deviation about the mean. 5hmC modifications were detected in all amplicons.

Genomic sequence upstream of *DNMT1* encodes a conserved mitochondrial targeting sequence (MTS)

Having confirmed that mtDNA contains 5mC residues (from which 5hmC is also presumably derived), we sought the enzyme(s) that may be responsible for these modifications. We started by asking whether any of the known catalytically active mammalian DNA methyltransferases might be targeted to mitochondria. Upon examination of the 5' UTR and 5' flanking genomic DNA upstream of the published transcriptional start sites for DNMT1, we identified additional sequence (101 codons in human and 63 codons in mouse) containing no termination codons, in frame with the ATG reported to be the primary translational start codon for DNMT1 (Figure 2-6) (93). In these 5' flanking sequences for both human and mouse DNMT1, we have labeled the published nuclear translational start codons as "ATG3" because there are two additional in-frame codons for methionine located upstream (Figure 2-6), each in a moderate context for ribosome binding (94).

The MitoProt II algorithm (http://ihg.gsf.de/ihg/mitoprot.html) (95) predicted isoforms for DNMT1 containing this additional N-terminal sequence in all mammalian species for which the 5' flanking sequence of DNMT1 was available. A protein produced from at least one of the putative upstream ATG codons was predicted with very high probability to translocate to the mitochondria (Table 2-4). The genome databases contained upstream sequences for human, mouse, chimpanzee, rat and cow *DNMT1*, and in each species, one or more in-frame potential start codons encoded a peptide with an extremely strong probability for mitochondrial localization (Table 2-4). Although the sequence conservation between them is low, all of these N-terminal extensions were predicted to form amphiphilic α-helices, a hallmark of other highly-characterized mitochondrial targeting sequences (Figure 2-6) (6, 7, 85).

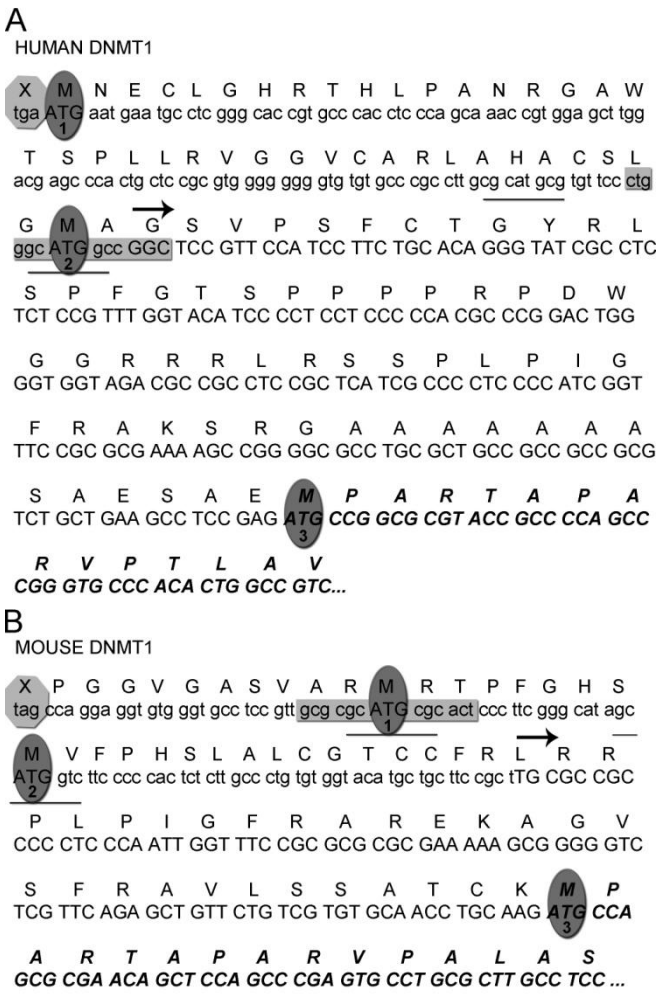


Figure 2-6: Genomic sequence upstream of DNMT1 encodes a MTS. (A) Sequence encoding an additional 101 in-frame amino acids, including two potential translation start ATG codons, is present upstream of the published human DNMT1 start ATG (ATG3). (B) In mouse, 63 in-frame amino acids are encoded upstream of the published translational start site, with two additional potential translation start ATGs. Overlapping binding sites for p53 (underlined) and NRF1 (gray box) are located across one of the two potential start ATG codons in both species. Published transcription start sites for DNMT1 are indicated by horizontal arrows above the sequence.

Table 2-4: Conservation of DNMT1 mitochondrial targeting sequence across mammalian species.

Species	ATG number	Position upstream of ATG*	Probability of Export to Mitochondria
Mouse	ATG1	+53aa	0.92
	ATG2	+45aa	0.96
	ATG3*	--	0.36
Human	ATG1	+101aa	0.84
	ATG2	+62aa	0.72
	ATG3*	--	0.48
Chimp	ATG1	+100aa	0.92
	ATG2	+62aa	0.70
	ATG3*	--	0.47
Rat	ATG1	+53aa	0.90
	ATG2*	--	0.33
Cow	ATG1	+97aa	0.43
	ATG2	+86aa	0.14
	ATG3	+67aa	0.14
	ATG4	+33aa	1.00
	ATG5*	--	0.97

ATG* = Published translational start site

The MTSs for human and mouse DNMT1 direct GFP to the mitochondria

We asked whether these upstream sequences directed an associated protein to a particular subcellular compartment. We amplified human and mouse mitochondrial targeting sequences, from ATG1 to just upstream of ATG3, and inserted them into the Vivid Colors pcDNA 6.2/C-EmGFP-GW/TOPO mammalian expression vector, in frame with a C-terminal GFP tag. These constructs were transfected into both NIH/3T3 mouse fibroblasts and HCT116 human colon carcinoma cells. Confocal microscopy demonstrated clearly that both human and mouse leader sequences targeted GFP to the mitochondria, as indicated by the distinct colocalization of MitoTracker Red stain with GFP fluorescence (Figure 2-7). Stained mitochondria from untransfected cells within the same visual field remained red in the merged images, serving as a negative control for GFP expression. Additionally, a chloramphenicol acetyltranferase (CAT)-GFP control plasmid remained cytosolic, and served as a negative control for nonspecific mitochondrial localization. We observed a similar result in HCT116 cells. The ability of both the human and mouse GFP-tagged constructs to operate across species indicates their functional conservation.

HCT116 cells transfected with the Vivid Colors pcDNA 6.2/C-EmGFP-GW/TOPO mammalian expression vector containing the cloned MTSs for human or mouse DNMT1 (the same constructs used for the confocal microscopy studies, Figure 2-7) were fractionated, separated on SDS-PAGE gels, and visualized by immunoblot. An antibody directed against the GFP-tag showed the presence of GFP in the mitochondrial fraction of cells transfected with either the human or mouse MTS, again demonstrating that these targeting sequences can operate across species (Figure 2-8). Interestingly, two different sets of protein bands were detected for GFP: the upper bands represent the intact translation products, carrying DNMT1 leader peptides

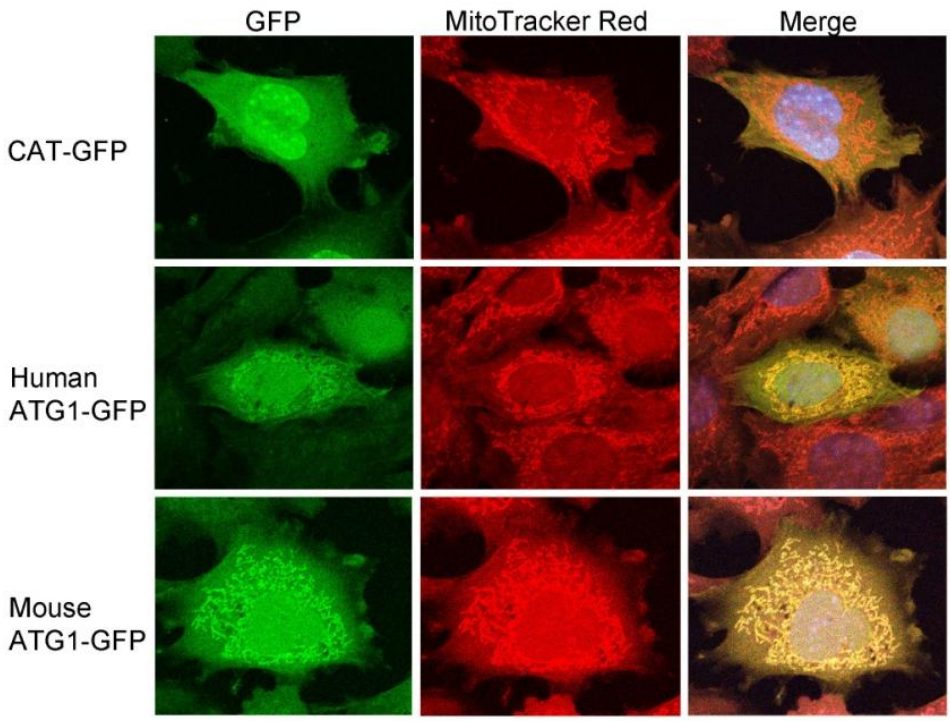


Figure 2-7: The MTS for DNMT1 directs GFP to the mitochondria. Transiently-transfected NIH/3T3 cells show distinct colocalization of GFP-tagged targeting sequences with MitoTracker Red stain in the mitochondria. A CAT-GFP negative control plasmid does not localize to the mitochondria. MTSs from both human and mouse operate across species, indicating their functional conservation. Confocal images taken by Dr. Shirley Taylor.

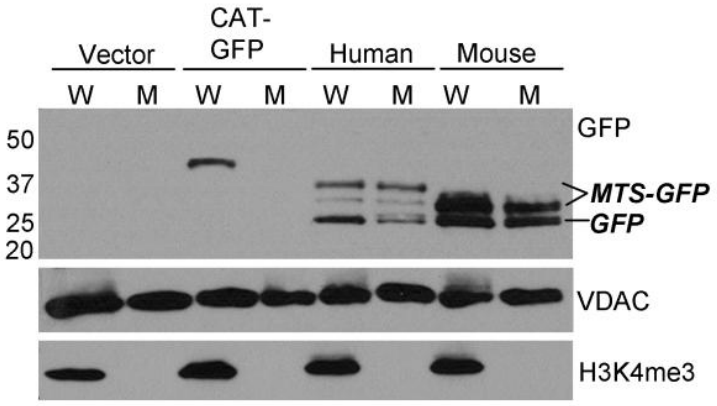


Figure 2-8: N-terminal peptides of mouse and human DNMT1 direct GFP to the mitochondria. Mouse and human N-terminal MTSs from ATG1 to immediately upstream of ATG3 were cloned in-frame to GFP and transiently transfected into HCT116 cells. Immunoblot analysis with an anti-GFP antibody shows the presence of GFP in the mitochondrial fraction (M). The upper bands represent intact translation product, carrying the DNMT1 MTS fused to GFP (MTS-GFP), whereas the lower bands are the expected molecular mass of GFP after proteolytic cleavage of the targeting peptides. W, whole cell lysate; M, mitochondrial lysate; MTS, mitochondrial targeting sequence; VDAC, voltage-dependent anion channel.

fused to GFP, whereas the lower bands are the expected molecular mass of GFP after proteolytic cleavage of the targeting peptides. Cells transfected with vector-only and CAT-GFP negative controls were free of GFP in the mitochondrial fraction, indicating that translocation of GFP to the mitochondria was specific to constructs containing a functional MTS. An antibody to VDAC was used here as both a mitochondrial marker and a loading control, and an antibody to H3K4me3 was used as a nuclear marker. This result, in conjunction with the immunoblots of subcellular fractions from human and mouse cells (Figure 2-10), suggests that mtDNMT1, much like many other nuclear-encoded mitochondrial proteins (6, 7, 85), undergoes proteolytic processing upon entry into the mitochondrial compartment.

RT-qPCR detects mature transcripts capable of encoding MTSs in human and mouse

Using sense primers positioned over ATG1 (mouse) or ATG2 (human) and anti-sense primers that overlap the exon 1-2 boundary by four nucleotides, endpoint RT-PCR detected mature transcripts capable of encoding the N-terminal mitochondrial targeting peptide in both human and mouse cell lines (Figure 2-9). No reverse transcriptase ("No RT") cDNA templates served as negative controls for PCR. These findings further supported the utilization of an upstream transcription start site to generate a mitochondrial-specific isoform of DNMT1.

Immunoblots of subcellular fractions from human and mouse cell lines demonstrate the presence of DNMT1, but not DNMT3a or DNMT3b, in the mitochondria

Immunoblots of subcellular fractions from mouse embryonic fibroblasts (MEFs) and HCT116 human colon carcinoma cells showed the presence of DNMT1 in the highly purified mitochondrial fraction (Figure 2-10). Neither DNMT3a nor DNMT3b could be detected. Both

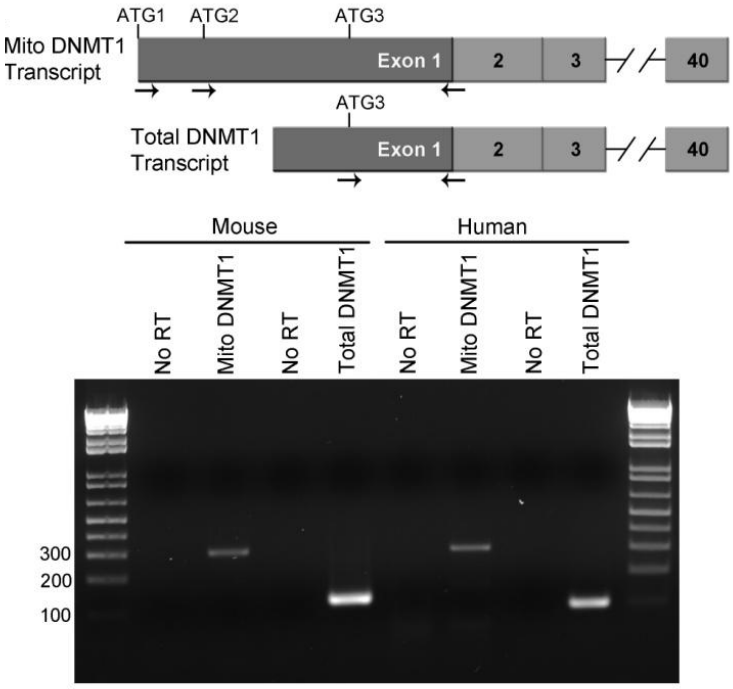


Figure 2-9: Detection of mature transcripts for mtDNMT1 in human and mouse. Mature transcripts capable of encoding the upstream MTS are detected in both human and mouse cells. No RT, no reverse transcriptase negative control.

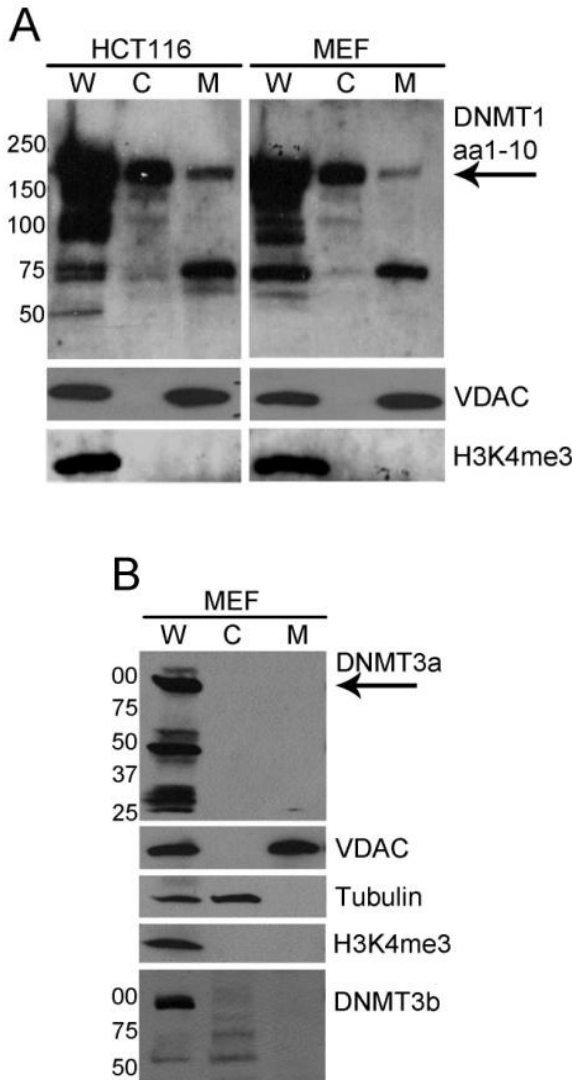


Figure 2-10: Detection of two protein species for DNMT1, and absence of the *de novo* DNMTs (DNMT3a and DNMT3b) in mitochondria. (A) A DNMT1 antibody directed against an N-terminal epitope (amino acids 1-10) detects two protein species in the mitochondrial fraction, full-length DNMT1 (185 kDa, arrow) and a 60 kDa peptide. Immunoblots performed by Erica Peterson. (B) DNMT3a or DNMT3b *de novo* DNA methyltransferases are not detected in mitochondria. Antibodies against compartment-specific markers: voltage-dependent anion channel (VDAC) (mitochondrial), tubulin (cytosolic) and H3K4me3 (nuclear) demonstrate the purity of each subcellular fraction. W, whole cell lysate; C, cytosolic lysate; M, mitochondrial lysate. For DNMT3b, the mitochondrial fraction was overloaded two-fold to exclude the possibility of low levels of DNMT3b in the mitochondria.

the full-length DNMT1 and a smaller peptide (~60kDa) are detected using an antibody directed against an N-terminal epitope of DNMT1, suggesting that proteolytic processing of mtDNMT1 occurs upon entry into the mitochondrial compartment. Antibodies against compartment-specific markers demonstrate the purity of each subcellular fraction. Voltage-dependent anion channel (VDAC) is an outer mitochondrial membrane protein; an antibody against VDAC was used as a mitochondrial marker. Histone H3 Lysine 4 trimethyl (H3K4me3) is a modified histone protein, found only in the nucleus, so the H3K4me3 antibody was used as a nuclear marker. An antibody to tubulin, which is a cytosolic protein, was used as a cytosolic marker. Absence of the nuclear marker, H3K4me3, in the mitochondrial fraction indicated purity of this fraction from contamination by nuclear material, which is the primary site of localization of DNA methyltransferases.

mtDNMT1 expression is regulated by factors that respond to oxidative stress

MatInspector software (http://www.genomatix.de/en/index.html) predicted a binding site for nuclear respiratory factor 1 (NRF1) within the MTS of both human and mouse *DNMT1* (Figure 2-6). This consensus sequence was located over one of the upstream in-frame start codons and was conserved in all other mammalian species examined for the presence of a MTS (Table 2-5). Under conditions of oxidative stress, NRF1 works in conjunction with PGC1α, which has been termed the "master regulator" of the response to oxidative stress, to up-regulate multiple nuclear-encoded mitochondrial proteins (1, 3). Accordingly, we transiently transfected NRF1, PGC1α, or both together, into HCT116 cells and fractionated transfected cells to analyze changes in the levels of mtDNMT1 by immunoblot. A moderate increase in mtDNMT1 expression was seen in cells transfected with NRF1 or PGC1α alone, whereas cotransfection

with both NRF1 and PGC1α resulted in a much more substantial increase in mtDNMT1 expression, relative to control (Figure 2-11). These results demonstrate that mtDNMT1 is a *bonafide* nuclear-encoded mitochondrial protein, sensitive to regulation by transcriptional activators that respond to oxidative stress.

Loss of p53 preferentially increases mtDNMT1 expression

Early studies performed by Erica Peterson in our laboratory identified a consensus p53 binding site that, when bound by p53, serves to repress DNMT1 transcription (Figure 2-6, Table 2-1) (89). This study reported a 3- to 6-fold increase in DNMT1 transcription following either activation or genetic deletion of p53 in both HCT116 cells and MEFs (89). Because p53 is known to regulate mitochondrial respiration (88), we questioned whether this tumor suppressor protein affected mRNA expression of the mitochondrial DNMT1 isoform. Using RT-qPCR with primers that distinguish the mitochondrial DNMT1 transcript from the total DNMT1 transcript (Figure 2-9), we measured the relative abundance of mtDNMT1 and total DNMT1 in both WT and p53-/- MEFs. The mitochondrial transcript comprises only 1-2% of the total DNMT1 mRNA synthesized in log-phase WT MEFs. However, upon loss of p53, the relative abundance of the mtDNMT1 transcript increased 6-fold, whereas the total DNMT1 mRNA increased only 3-fold. This suggests a preferential up-regulation of the mitochondrial transcript in cells lacking p53 (Figure 2-12A).

Immunoblot analysis of subcellular fractions from both WT and p53-/- MEFs recapitulated this phenomenon at the protein level (Figure 2-12B). Whole cell, cytosolic and mitochondrial fractions from these isogenic cell lines were separated by SDS-PAGE, and probed with antibodies directed against DNMT1 and appropriate compartment-specific controls. The

Table 2-5: Consensus sequences for putative NRF1 binding sites within the MTS for DNMT1.

NRF1 consensus	YGCGCAYGCGCR	Base Pair Position
Mouse	CGCGCATGCGCA	-164 to -153
Human	TGGGCATGGCCG	-191 to -180
Chimp	TGGGCATGGCCG	-191 to -180
Rat	TGCGCATGCGCA	-164 to -153
Cow	CGCGCATGCGCG	-205 to -194

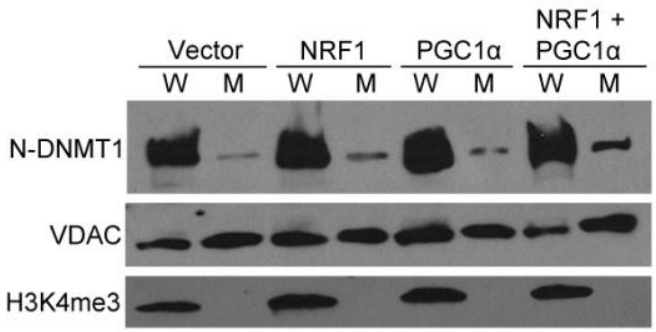


Figure 2-11: Regulation of mtDNMT1 expression by NRF1 and PGC1α. HCT116 cells were transiently transfected with recombinant NRF1 and PGC1α constructs and fractionated into whole cell and mitochondrial fractions. Immunoblot analysis with an anti-DNMT1 antibody shows that either NRF1 or PGC1α alone induce a modest up-regulation of DNMT1 in the mitochondrial fraction. However, cells co-transfected with both NRF1 and PGC1α demonstrated a substantial increase in mtDNMT1 expression. W, whole cell lysate; M, mitochondrial lysate.

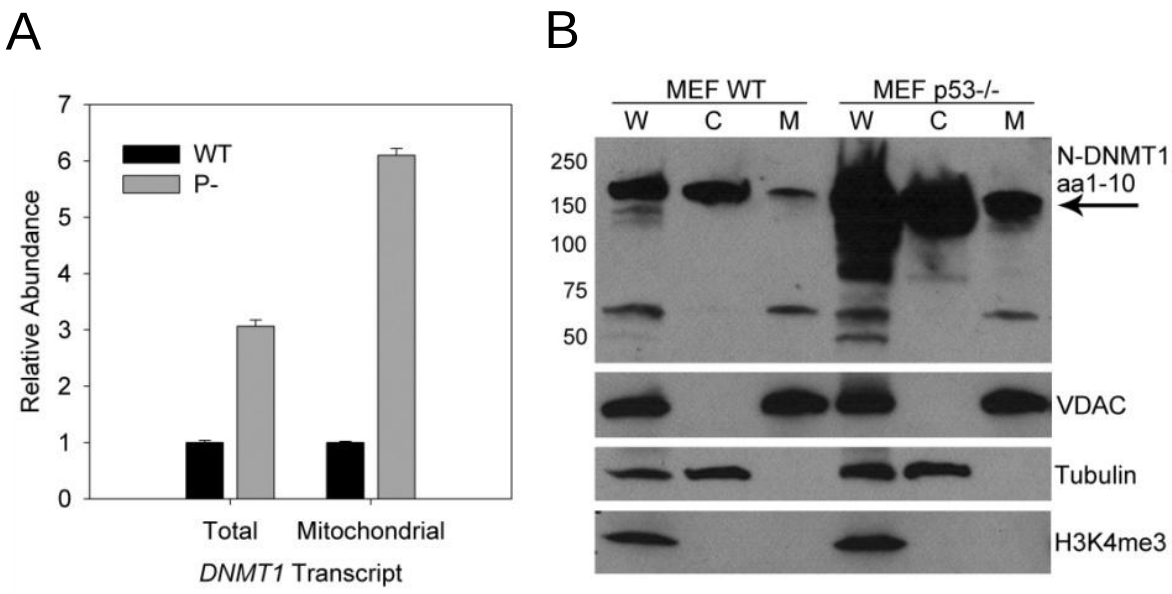


Figure 2-12: Regulation of mtDNMT1 expression by p53. (A) Loss of p53 selectively up-regulates mtDNMT1 expression. Values represent mean +/- SD for two independent experiments. WT, wild-type MEFs; P-, p53-null MEFs. (B) Immunoblot analysis of WT and p53-/- MEFs shows a striking increase mtDNMT1 expression in p53-null cells. W, whole cell lysate; C, cytosolic lysate; M, mitochondrial lysate.

blots illustrate a striking increase in full-length DNMT1 expression in the mitochondrial fraction of cells lacking p53. Interestingly, no change in the level of the 60kDa band was observed, suggesting that this immunoreactive protein is either an artifact due to non-specific binding of the antibody, or that cleavage of the full-length protein into the 60kDa species is limited by the mitochondrial processing peptidase.

Up-regulation of mtDNMT1 is associated with gene-specific changes in mitochondrial transcription

Using the same WT and p53-null isogenic MEF cell lines, we asked whether this drastic overexpression of mtDNMT1 might be reflected in an alteration of mitochondrial transcription. Prashant Thakkar in our laboratory used random hexamer-primed cDNA and RT-qPCR with various primers specific to regions along the mitochondrial genome, and we observed an interesting contrast of effects (Figure 2-13). Expression of NADH dehydrogenase subunit 6 (ND6), the only protein-coding gene on the light strand of mtDNA (Figure 1-2), was significantly reduced in response to overexpression of mtDNMT1. This suggests a role for methylation of mtDNA in repression of light-strand transcription, an effect that would mimic the role DNA methylation plays in the nucleus. On the heavy strand, ATPase subunit 6 (ATP6) and cytochrome *c* oxidase subunit 1 (COX1) were unaltered in their expression levels. However, expression of NADH dehydrogenase subunit 1 (ND1), the first protein-coding region on the heavy strand, immediately following the ribosomal RNA (rRNA) genes (Figure 1-2), was significantly increased in response to elevated mtDNMT1. This increase in ND1 expression, without changes in expression of ATP6 or COX1 located further downstream on the heavy strand, suggests a gene-specific effect on mitochondrial transcription.

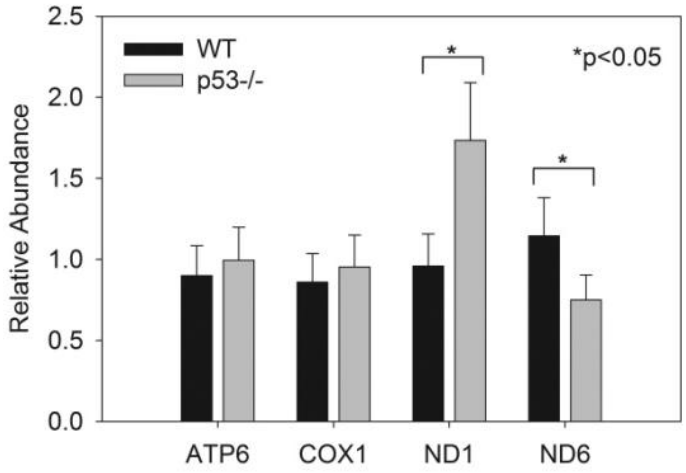


Figure 2-13: Loss of p53/up-regulation of mtDNMT1 induces gene-specific effects on mitochondrial transcription. Using random-hexamer-primed cDNA, gene-specific changes in mitochondrial transcription were observed upon over-expression of mtDNMT1, relative to nuclear 18S rRNA. To test for differences in expression of four mitochondrial genes across the two cell lines, we applied an ANOVA using least-square means. ND1 and ND6 showed significant differences in expression ($P < 0.05$). In p53-/- MEFs, ND1 mRNA expression was significantly higher (mean = 1.73, SE = 0.18, 95% CI = 1.37, 2.09) than in WT MEFs (mean = 0.95, SE = 0.09, 95% CI = 0.76, 1.15). ND6 mRNA expression in p53-/- MEFs was significantly lower (mean = 0.75, SE = 0.07, 95% CI = 0.59, 0.90) than in WT MEFs (mean = 1.14, SE = 0.11, 95% CI = 0.91, 1.38). Studies performed by Prashant Thakkar.

To be sure that we were not masking any transcriptional effects by looking at the average across both H and L strands of mtDNA, we modified our cDNA synthesis strategy to generate only strand-specific-primed cDNA. The ATP6 gene was chosen as the region to initiate strand-specific cDNA synthesis because it was approximately equidistant from either region of interest on the H (ND1) and L (ND6) strands. Using a sense primer within the ATP6 gene, we synthesized only the light strand of mitochondrial cDNA; likewise, using an antisense primer within the ATP6 gene, we synthesized only the heavy strand. Total RNA from WT and p53-/- isogenic MEFs was used as the template for cDNA synthesis. RT-qPCR was performed using primers for heavy-strand genes against heavy-strand template cDNA and primers for light-strand genes against only light-strand template cDNA. Two new light strand primer sets were designed against non-coding sequence complementary to ND4 (non-coding primer set "NC1") and ND5 (non-coding primer set "NC2") on the heavy strand, in order to get a more complete picture of changes in transcription along the light strand (Table 2-2). Increased expression of mtDNMT1 appeared to differentially affect transcription initiating from HSP1 and HSP2. Transcription of the ribosomal RNA genes, 12S and 16S rRNA, initiates from HSP1 on the heavy strand, and showed contrasting effects (Figure 2-14A). No change in 12S rRNA expression was observed between WT and p53-/- MEFs. However, a substantial decrease in expression of 16S rRNA was observed upon loss of p53/up-regulation of mtDNMT1 (Figure 2-14A). Heavy strand transcripts initiated at HSP2 all showed the same trend: expression of both ND1 and COX1 were significantly up-regulated by overexpression of mtDNMT1 in p53-/- MEFs (Figure 2-14B). This increase in ND1 is consistent with what we observed before using random hexamer-primed cDNA, looking at an average of transcription from both strands (Figure 2-13). However, using a modified cDNA synthesis strategy, we were able to see an increase in COX1 expression, which

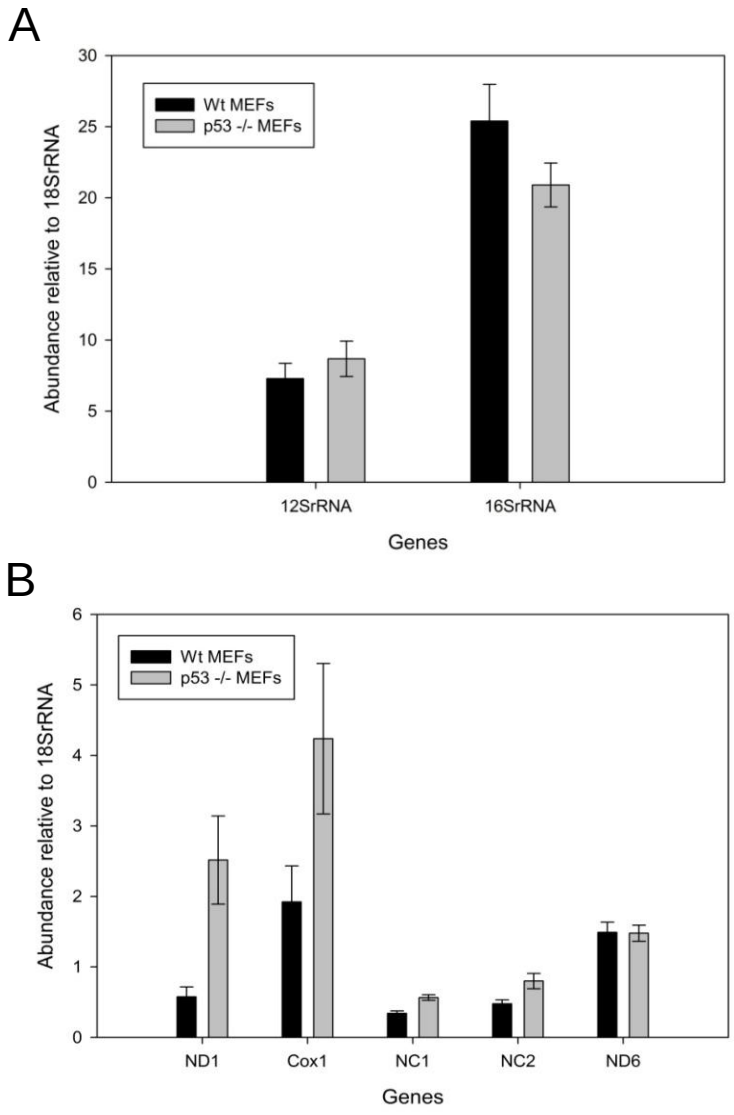


Figure 2-14: Loss of p53/up-regulation of mtDNMT1 induces additional gene-specific effects on mitochondrial transcription. Total RNA from WT and p53-/- MEFs was used to generate either random hexamer-primed cDNA (for nuclear 18S rRNA control) or gene-specific-primed cDNA with primers located within the ATP6 gene. Gene-specific changes in mitochondrial transcription are again observed upon over-expression of mtDNMT1, relative to nuclear 18S rRNA. NC1, non-coding region complementary to ND4 on the H strand; NC2, non-coding region complementary to ND5 on the H strand. Studies performed in collaboration with Prashant Thakkar.

was somehow masked before, by looking at the average across both strands. On the light strand, looking at transcripts initiated at the light strand promoter (LSP), we saw increases in both NC1 and NC2 mRNA in response to mtDNMT1 overexpression (Figure 2-14). Interestingly, looking only at the light strand, there was no change in ND6 transcription upon overexpression of mtDNMT1 (Figure 2-14), which was not consistent with what we observed before (Figure 2-13).

DISCUSSION

Previous studies likely underestimated the level of epigenetic modification within the mitochondrial genome

Cytosine methylation of the mitochondrial genome has remained largely overlooked, in part because early reports using nearest-neighbor analysis indicated that this modification was present at only 3-5% of CpG dinucleotides (17). This is well below the level of methylation (~60-80% of CpG dinucleotides) seen in the nucleus (96, 97). The data presented here show a 10- to 20-fold enrichment of mtDNA sequences in immunoprecipitates using a 5mC antibody (Figure 2-3A). This level of enrichment is somewhat lower than that usually obtained from total gDNA (~100-fold for CpG islands), and likely reflects the CpG-sparse nature of the mitochondrial genome, which does not contain CpG islands (17, 18). Additionally, our data demonstrates the presence of 5hmC residues in mtDNA using two independent assays (Figures 2-3B, 2-4 and 2-5). Our data shows a 38-580-fold enrichment of mtDNA sequence in immunoprecipitates using a 5hmC antibody, which is in striking contrast to the low level of background signal in our nonspecific IgG controls (Figure 2-3A). Although we cannot use the MeDIP assay with two antibodies of different specificities to accurately quantitate the frequency of each type of modified residue, the β-glucosyltransferase assay is quantitative. However, this

assay relies upon enzymatic digestion by *Gla*I, so the detection of 5hmC residues is only possible at specific *Gla*I cleavage sites. While the β-gt assay can provide valuable information as to the relative abundance of 5hmC residues within sequences containing a given number of *Gla*I restriction sites (Figure 2-5), these sites are not representative of the entire mitochondrial genome. Regardless, our findings that both 5mC and 5hmC residues exist within mtDNA suggest that epigenetic modification of cytosines in the mitochondrial genome is likely to occur more frequently than previously thought.

Profiles of 5mC- and 5hmC-modified residues across the mitochondrial genome

Mapping the location and distribution of 5mC or 5hmC residues in either the nuclear or mitochondrial genomes is not yet technically feasible because the modified bases are indistinguishable by bisulfite modification and sequencing (56, 98). However, MeDIP analysis of specific DNA sequences, using antibodies specific for either modification can distinguish between 5mC and 5hmC residues. While we cannot compare the level of enrichment of mtDNA sequences we see with the 5mC antibody to that with the 5hmC antibody, we can compare across regions immunoprecipitated with the same antibody. Our data illustrates that there are easily detectable levels of both 5mC and 5hmC across all tested regions of the mitochondrial genome.

Primer set 27 is a particularly interesting case, as it overlaps the D-loop control region, which contains the promoters responsible for initiating mtDNA transcription (Figures 1-2 and 2-3) (1, 3). The D-loop exists as a stable triple-helical structure, containing an RNA primer that is required for initiation of mtDNA replication (1). We have found this region to be resistant to *in vitro* methylation by *M.Sss*I, a bacterial cytosine methyltransferase. It is therefore possible that the kinetics of epigenetic modification in this region of the mitochondrial genome may be

different from those in coding regions. For instance, the complex structure of the D-loop may prevent access of mtDNMT1 to the DNA, or if mtDNMT1 can gain access, perhaps base-flipping of the target cytosine residue out of the DNA helix is prevented by the stability of the triplex D-loop structure. Either of these hindrances would limit modification of cytosines within the D-loop to 5mC. Additionally, if 5mC residues are required as precursors for 5hmC, then limiting 5mC modification would also negatively impact the level of 5hmC in this region. Nonetheless, both 5mC and 5hmC modifications are present in the amplicon containing the D-loop, albeit, perhaps to a lesser extent than other regions across the mitochondrial genome.

Possible mechanisms for generation of 5hmC

In the nucleus, 5hmC is generated from 5mC through the action of the ten-eleven translocation (TET) family of methylcytosine oxygenases (46). There is not yet concrete evidence regarding the presence or absence of these enzymes in mitochondria. We therefore cannot rule out the possibility of a different mechanism for the generation of 5hmC in mtDNA, including the covalent addition of 5-hydroxymethyl groups directly to DNA cytosine residues by mtDNMT1, as described by Liutkeviciute and colleagues (Figure 2-15) (48). This mechanism would require formaldehyde generated by mitochondrial mixed-function oxidases, which is usually removed by aldehyde dehydrogenases as a protective measure against formaldehyde-induced oxidative damage (99). This uncharacteristic reaction would also require the base-flipping activity of DNMT1 to perform nucleophilic addition of an exogenous aldehyde to the C5 position of a target cytosine, rather than its naturally-catalyzed nucleophilic substitution reaction (Figure 2-18). Using aldehydes, which lack an adenosyl anchor moiety and therefore are not true "cofactors", would be an atypical chemical reaction for DNMT1, which is normally considered

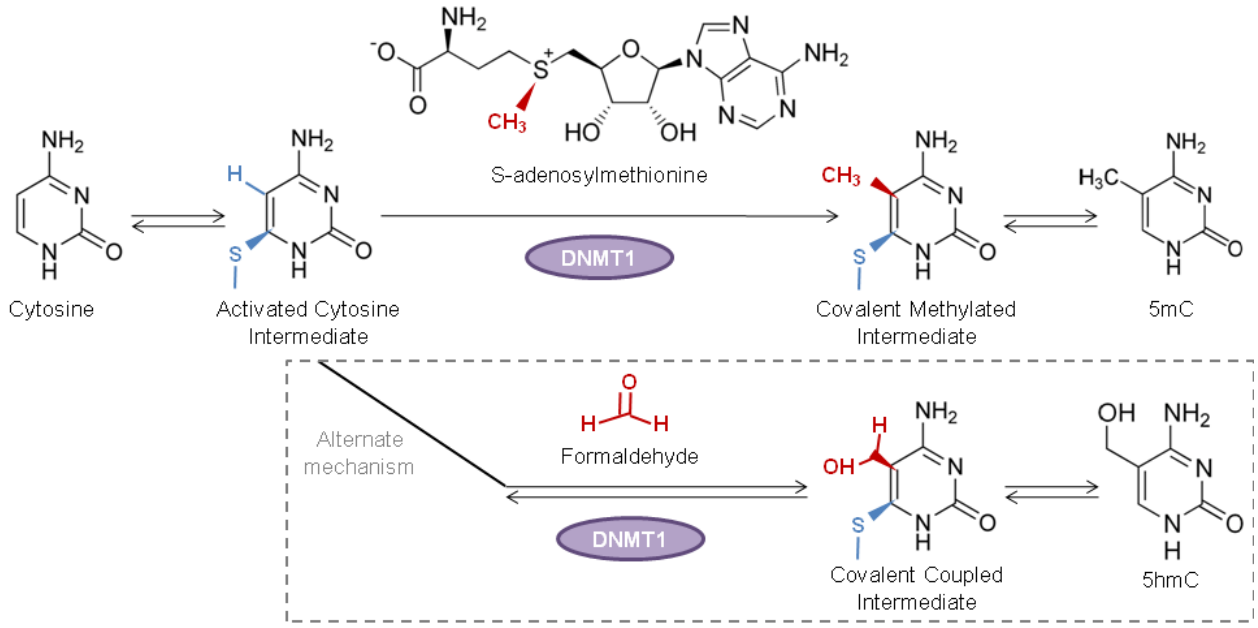


Figure 2-15: Mechanism for direct addition of a 5-hydroxymethyl group to cytosine by DNMT1. Methylation by C-5 methyltransferases occurs via a direct nucleophilic S_N2 attack of the activated cytosine intermediate onto the cofactor, S-adenosylmethionine. This yields the covalent methylated intermediate, and then 5mC. In an alternate mechanism, the activated cytosine intermediate can undergo a nucleophilic addition reaction with an exogenous aldehyde. This gives rise to a covalent coupled intermediate, and ultimately releases the 5hmC product. Adapted from figure by Liutkeviciute, *et al*. (2009) *Nat Chem Biol* 5: 400-402.

to be a cofactor-dependent enzyme (48). However, the mitochondrial matrix is a unique environment, and may provide conditions that allow this reaction to proceed.

Alternatively, 5hmC may arise in a mitochondrial folate-dependent manner, as a modification of dCMP by an enzyme analogous to that found in T-even bacteriophage. In bacteriophage, the folate-dependent dCMP hydroxymethylase catalyzes the formation of 5hm-dCMP from dCMP (100). 5hm-dCMP can be converted to 5hm-dCTP through the action of deoxycytidylate kinases, and subsequently incorporated into bacteriophage DNA. Incorporated 5hmC residues can then be further modified by β-glucosyltransferase enzymes, in a reaction identical to the one exploited in the β-glucosltransferase assay for sequence-specific detection of 5hmC in mtDNA. These mechanisms exist in T-even bacteriophage as a protective measure, to ensure that phage DNA is not subjected to host-mediated restriction endonuclease cleavage. For this to represent a viable mechanism for generation of 5hmC in mitochondria, a mitochondrial enzyme with activity analogous to phage dCMP hydroxymethylase would be required, and would not rely on the prior formation of 5mC.

Functional significance of mtDNA epigenetic modification

a. What are the possible functions of 5mC and 5hmC in mtDNA?

Overexpression of mtDNMT1, as seen in our p53-/- MEFs (Figure 2-12), results in an asymmetric, gene-specific alteration in mitochondrial transcription patterns, suggesting diverse roles for mtDNMT1 and cytosine modification in this organelle. Looking first at the results obtained using strand-specific-primed cDNA (Figure 2-14), we observed decreased expression of 16S rRNA on the heavy strand when mtDNMT1 was overexpressed. This effect mimics what we might expect to observe in the nucleus, where methylation status almost always inversely

correlates with gene expression. However, increased expression of ND1 and COX1 on the heavy strand, and NC1 and NC2 on the light strand suggest an opposite mode of action for methylation of mtDNA. It could be that methylation of mtDNA serves to recruit the mitochondrial transcription machinery, enhancing expression of methylated genes, rather than repressing them as methylation does in the nucleus. Alternately, methylation could influence the structure of mtDNA, such that either the methylation signal itself or binding by mitochondrial 5mC-specific binding proteins opens the supercoiled configuration of mtDNA, inducing a more transcriptionally-active state, much like acetylation promotes a more euchromatic state of nuclear DNA.

Increases in transcription of ND1 and COX1 initiating at HSP2, with a concurrent decrease in transcription of 16S rRNA from HSP1 suggests that the methylation-mediated effect is rooted in transcription initiation. Perhaps mtDNA methylation differentially affects transcripts originating at either H-strand promoter, providing an additional level of control of rRNA transcript production relative to protein-coding mRNA in mitochondria.

However, random-hexamer-primed cDNA (Figure 2-13) shows no change in expression of ATP6 or COX1 on the H-strand, but a significant increase in transcription of ND1, which is consistent with the results obtained using strand-specific-primed cDNA. Transcription of all three of these protein-coding genes initiates at HSP2, suggesting a methylation-mediated effect that is distinct from transcription initiation. Instead, mtDNA methylation may interfere with proper termination of transcription of the rRNA genes initiating at HSP1, resulting in increased readthrough from HSP1 to the next open reading frame, ND1. A schematic representation of this proposed effect is depicted in Figure 2-16, and is discussed in more detail in the next section.

The function of 5hmC in the nuclear genome is not yet clear. It has been proposed that 5hmC is an intermediate metabolite in either active demethylation of the genome by repair enzymes (101, 102), or passive demethylation as a result of lack of recognition by enzymes involved in maintenance methylation (101, 103). Evidence for the former hypothesis is stronger, as a recent study described a TET1-mediated oxidation-deamination mechanism for active DNA demethylation in the brain (47). It has also been proposed that 5hmC alters local chromatin structure because 5hmC is not recognized by 5mC-binding proteins (101). If mtDNA methylation or binding by 5mC-specific binding proteins serves to open the configuration of mtDNA as discussed above, then perhaps 5hmC modifications induce a more compact mtDNA configuration, resembling a "heterochromatic" state, and repressing mitochondrial transcription.

Studies using ChIP-seq to map the positions of the TET1 enzyme bound to nuclear DNA demonstrated that TET1 was bound to CpG-rich promoters (101), contrasting the trends seen for nuclear methylation, where 5mC is normally absent at CpG-rich promoters. Furthermore, it has been shown that DNMT1 does not bind to its target cytosine when positioned opposite from a 5hmC residue in hemi-hydroxymethylated DNA (103). Together, these findings suggest that 5hmC may serve to prevent inappropriate methylation and silencing of CpG-rich genes. Thus, the roles for 5mC and 5hmC in mitochondria could be reversed from their functions in the nucleus. Nonetheless, if 5mC represents methylation and 5hmC represents demethylation of mtDNA, then the levels of these marks in mitochondria are likely in a delicate balance, affected themselves by enzymatic function, and with downstream consequences including changes in mitochondrial gene expression.

b. Proposed role for mtDNMT1 in mitochondria

Although DNMT1 is generally considered to be the maintenance DNA methyltransferase, it is able to methylate completely unmethylated DNA *in vitro* with an efficiency that exceeds that of the *de novo* methyltransferases, DNMT3a and DNMT3b (59). Thus, DNMT1 appears to be capable of both initiating and maintaining cytosine methylation in the nucleus, and the lack of *de novo* methyltransferases in mitochondria implies that mtDNMT1 is involved in both processes in this organelle. The presence of modified cytosines in the D-loop control region of mtDNA, which houses the H-strand origin of replication and the promoters driving transcription initiation of both heavy and light strands, supports a role for mtDNMT1 in the regulation of mitochondrial gene expression.

We observed a significant increase in ND1 expression, regardless of whether we looked at an average of both heavy and light strand messages (Figure 2-13), or the heavy strand only (Figure 2-14). We hypothesize that either mtDNMT1 itself or methylation by mtDNMT1 may be interfering with proper termination of transcription from HSP1 (Figure 2-16). A binding site for mitochondrial termination factor 1 (mTERF1) is located between the end of the 16S rRNA gene and the translation start of ND1 (Figure 1-2) (1). mTERF1 binds to both HSP1 and the terminator binding site, forming a transcription loop that maintains high-level production of rRNA. Transcripts initiating at HSP2 produce polycistronic messages encoding the entire heavy strand (1). Our data raise the possibility that mtDNMT1, either through modification of CpG dinucleotides or by direct protein-protein interaction, interferes with mTERF-dependent transcription termination, allowing read-through from HSP1 to the next transcriptional unit (ND1) without impacting polycistronic mRNA synthesis from HSP2 (Figure 2-16).

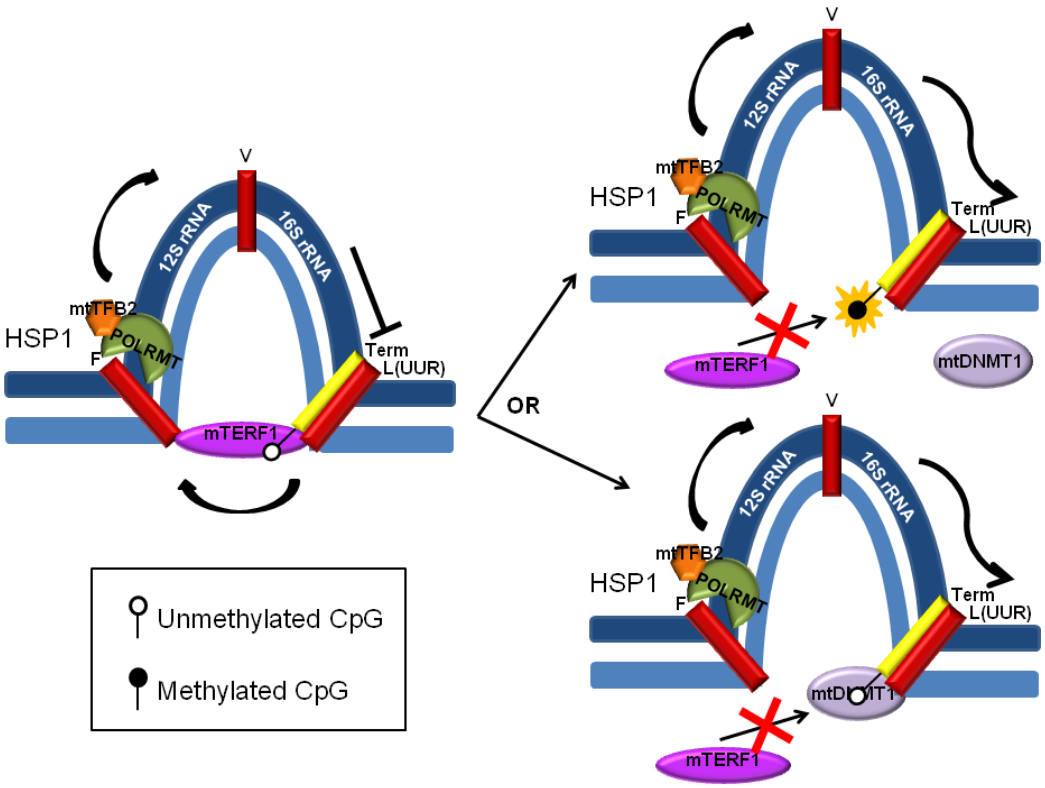


Figure 2-16: Proposed mechanism for up-regulation of ND1 expression. Normally, the transcription terminator, mTERF1 binds simultaneously to HSP1 and its Term site (yellow bar), forming a rDNA loop to yield high level transcription of the rRNA genes. When mtDNMT1 is up-regulated, we observe an increase in transcription of ND1, the first protein-coding gene on the H strand, immediately downstream of the rRNA genes and Term site. We hypothesize that either: 1) methylation of a CpG dinucleotide within the Term site, or 2) direct binding of mtDNMT1 at the Term site, serves to disrupt proper termination of transcription from HSP1. Failure to terminate H strand transcription at the Term site would result in readthrough of the transcriptional machinery to the next coding unit (ND1), causing increased expression of ND1, and possibly other ORFs downstream of ND1.

c. Are the observed effects on mitochondrial transcription gene-specific or strand-specific?

Results from the first RT-qPCR experiments, which measured changes in mitochondrial gene expression in isogenic WT and p53-/- MEFs using random hexamer-primed cDNA (Figure 2-13), appeared to suggest that the effects of overexpression of mtDNMT1 may be strand-specific, as we observed an increase in ND1 transcription on the heavy strand with a decrease in ND6 transcription on the light strand. However, the absence of change in ATP6 or COX1 expression, both on the heavy strand, implied that the effect of mtDNMT1 overexpression on mitochondrial transcription was more complicated than anticipated, and led us to focus our attention on transcription of one mtDNA strand at a time.

RT-qPCR measuring changes in mitochondrial gene expression in WT and p53-/- MEFs, this time from strand-specific primed cDNA (Figure 2-14), also suggested a gene-specific phenomenon was at play. The increase in ND1 expression was reiterated using this modified strategy, so this effect is consistent between the two experiments. However, an increase in COX1 expression (heavy strand) was also revealed using strand-specific-primed cDNA, an effect that we predict may have been hidden using random-hexamer-primed cDNA and looking at an average across both strands.

One major difference in the results obtained from the two different cDNA priming strategies was in the expression of ND6: using random-hexamer-primed cDNA, we observed a statistically significant decrease upon upregulation of mtDNMT1 (Figure 2-13), but this difference disappeared using strand-specific-primed cDNA (Figure 2-14). This apparent "loss" of an effect is likely a technical artifact, resulting from the method used to generate cDNA and

differences in priming efficiencies between the ATP6asn and ATP6sn primers used to synthesize the H- and L-strands, respectively.

Other changes in mitochondrial gene expression were also observed upon mtDNMT1 overexpression using strand-specific-primed cDNA (Figure 2-14): a decrease in 16S rRNA (heavy strand), which is consistent with other reports in the literature of a p53-dependent effect on mitochondrial transcription (86, 104), and increases in NC1 and NC2 (light strand). Consistent changes in expression of genes on each strand would afford greater confidence in concluding that mtDNMT1 affects mitochondrial transcription in a strand-dependent manner. However, the variability of effects from gene to gene along the same strand suggest that overexpression of mtDNMT1 has a gene-specific effect on mitochondrial transcription.

d. Are the observed effects on mitochondrial transcription p53-dependent or mtDNMT1-dependent?

We hypothesize that the observed differential effects on mitochondrial transcription are mtDNMT1- or methylation-dependent. However, we cannot ignore the confounding effects due to altered p53 status, which is an obvious limitation of these experiments. The isogenic WT amd p53-/- MEF cell lines were used to induce changes in mtDNMT1 expression because they were readily available, but this is not a clean system. Therefore, at this time, we cannot conclude that p53 status is not contributing to these transcriptional effects. Other labs have reported that p53 influences mtDNA copy number, showing that cells lacking p53 exhibited a significant reduction in mtDNA content compared to WT cells (105, 106). If this is true, then we would expect a reduction in mtDNA content to be reflected by a generalized decrease in mitochondrial transcription. In our mitochondrial transcription experiments, we observe a decrease in 16S

rRNA; perhaps this effect can be explained by reduction in mtDNA copy number as a result of the loss of p53. However, the consistent increase in ND1 expression on the heavy strand, as well as increases in NC1 and NC2 on the light strand, does not conform to that theory. Such gene-specific effects are more likely a characteristic of gene-specific changes in mtDNA methylation profiles, or occupancy of specific sites along mtDNA by mtDNMT1. We propose to further investigate the effects of mtDNMT1 alone on mitochondrial gene expression by either over-expression or selective depletion of mtDNMT1. Some of these studies are described in more detail in Chapter 5.

Conservation of DNMT1 MTSs across mammalian species and thoughts on import and proteolytic processing of mtDNMT1

Our studies report a mitochondrial isoform of DNMT1, which is the only member of the catalytically active mammalian DNA methyltransferase family found in this organelle. The conservation of an ORF encoding a mitochondrial targeting sequence upstream of the commonly accepted translational start codon across multiple mammalian species suggests an important role for this enzyme in mitochondrial function.

Like the majority of other nuclear-encoded mitochondrial proteins, mtDNMT1 is directed to the organelle by way of a cleavable, N-terminal targeting sequence. Such presequences typically direct import of their mitochondrial precursor proteins to the mitochondrial matrix, employing translocase of the outer mitochondrial membrane (TOM), translocase of the inner mitochondrial membrane (TIM23) and presequence translocase-associated motor (PAM) machinery (6, 7). Upon import into the matrix, the targeting presequence is cleaved by the mitochondrial processing peptidase (MPP), to generate the final functional protein. Our data

suggest that a similar import mechanism is used for mtDNMT1, as a protein of the expected size for GFP after cleavage of the DNMT1-MTS appeared in the mitochondrial fraction of cells transfected with our MTS-GFP constructs. In these experiments, the heterologous protein, GFP was used as both a fluorescent and an immunoreactive marker, allowing us to track the subcellular localization of the presumed MTS. These experiments prompted us to devise a model for mtDNMT1 import and subsequent proteolytic processing within the mitochondrial compartment (Figure 2-17).

Immunoblots of subcellular fractions from both human and mouse cell lines show that full-length DNMT1 (185kDa) is easily detected using an antibody directed against a far N-terminal epitope (aa1-10, Abcam), along with a smaller immunoreactive protein species approximately 60kDa in size (Figure 2-10). Interestingly, upon genetic deletion of p53, a dramatic increase in expression of only full-length mtDNMT1 is observed; expression of the 60kDa band remains the same between WT and p53-/- MEFs (Figure 2-12B). This suggests that, if this smaller protein species is in fact a proteolytic cleavage product of the full-length enzyme, it is not subject to the same p53-dependent regulation as the full-length enzyme. Alternatively, the lack of up-regulation of the 60kDa species could mean that the mitochondrial processing peptidase (MPP) is saturated and cannot maintain proteolytic processing of higher levels of mtDNMT1. It is also possible that this immunoreactive protein represents non-specific binding, and is simply an artifact of the antibody used. Because this 60kDa protein is detected with an N-terminal antibody, it would represent an N-terminal fragment, lacking the catalytic domain and enzymatic active site, and thus would likely not be active on mtDNA. Such a fragment would still contain the MTS for translocation to the mitochondrial compartment, and DNA binding domains, which may allow this truncated protein to compete with the catalytically-active form of

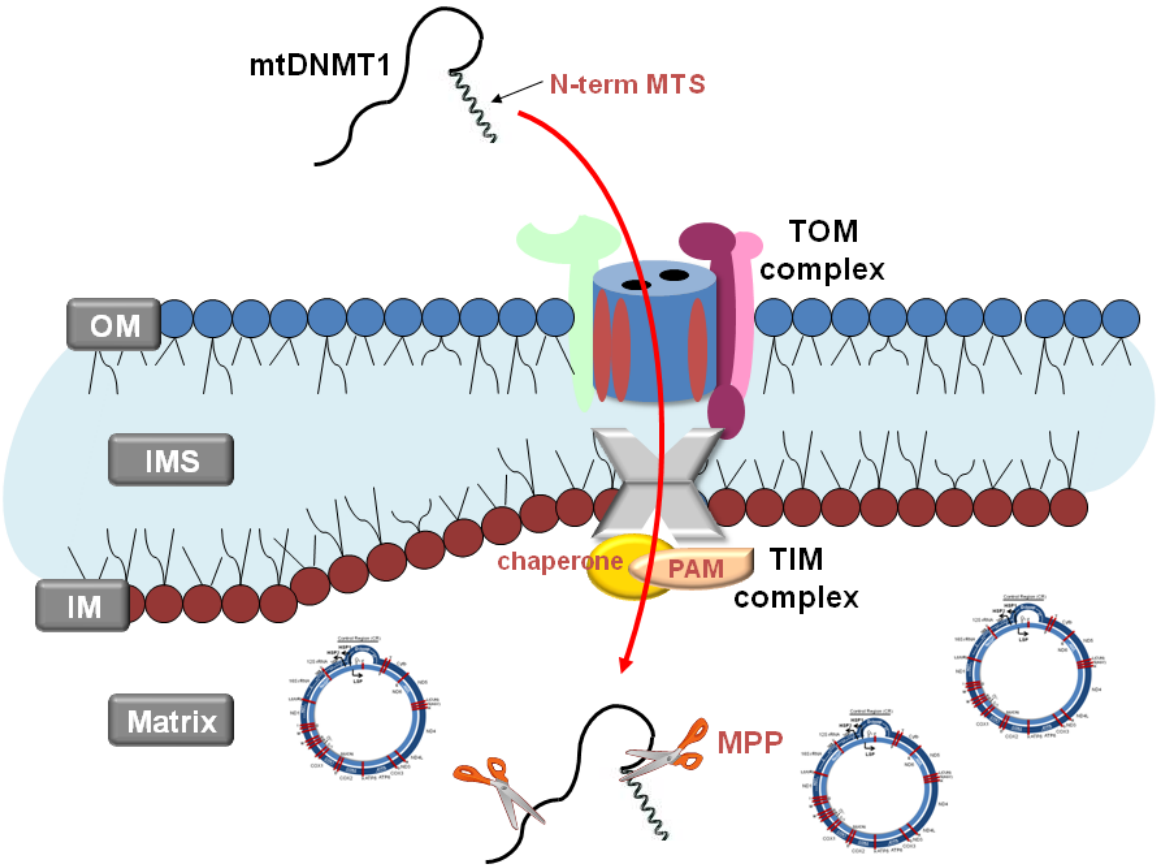


Figure 2-17: Model for translocation, import and proteolytic processing of mtDNMT1. This model is based on the most common route of import for nuclear-encoded mitochondrial precursor proteins containing an N-terminal mitochondrial targeting sequence (MTS). The red arrow indicates the flow of direction for import through the translocase machinery. The MTS for mtDNMT1 would first be recognized by receptors on the mitochondrial outer membrane (OM), which then direct its passage through the main entry gate for mitochondria, the translocase of the outer membrane (TOM) complex. At this stage, precursor proteins follow different sorting pathways, based on their ultimate function within the organelle. Since mtDNMT1 is expected to bind and methylate mtDNA, it would be directed to the matrix, which would involve traversing the inner membrane (IM) via the translocase of the inner membrane (TIM) machinery. The TIM complex is assisted by the presequence associated motor (PAM), and other specific chaperone proteins, to fully import the precursor protein into the mitochondrial matrix. Once inside the matrix, the precursor protein undergoes cleavage of its N-terminal MTS, and as we suspect for mtDNMT1, additional proteolytic processing events to generate the final functional form of mtDNMT1.

mtDNMT1 for binding to mtDNA, thereby imposing control over which sites become methylated.

Regulation of mtDNMT1 by factors controlling mitochondrial function

Conservation of not only the MTSs across species, but also the configuration of transcription factor binding sites within the MTSs across species (Tables 2-3, 2-4 and 2-5), is suggestive that the mtDNMT1 isoform has some evolutionary significance with regard to proper functioning of the organelle. It has been well established that NRF1 and PGC1α function as transcriptional coactivators that respond to conditions of oxidative stress by up-regulating expression of nuclear-encoded mitochondrial proteins, and stimulating mitochondrial biogenesis (1, 3). Our observations that mtDNMT1 is up-regulated in a manner similar to many other nuclear-encoded mitochondrial proteins (Figure 2-11) is evidence that mtDNMT1 is controlled by factors directly involved in mitochondrial function, and implicated mtDNA methylation in the mitochondrial response to cellular stress.

p53 is best known as the most frequently mutated gene in cancer, but it has also been shown to be intimately linked to mitochondrial function (86, 88, 104-106). Studies have reported that p53 is involved in regulating the balance between the utilization of respiratory and glycolytic pathways in the mitochondria by controlling expression of the subunit proteins forming the COX complex (88). Thus, regulation of the mtDNMT1 isoform by p53 (Figure 2-12) is suggestive that changes in expression or activity of mtDNMT1 may have downstream consequences affecting overall functioning of the organelle. Collectively, our studies on the regulation of mtDNMT1 demonstrate that this isoform is controlled by factors intimately linked with mitochondrial function, and lend further support to our hypothesis that mtDNMT1 is functionally relevant.

Is there a connection between the endosymbiotic theory and conservation of the catalytic domain of DNMTs across eukaryotes and prokaryotes?

Common to all of the DNMTs is the structure and amino acid composition of the catalytic domain (35, 59, 71). This C-terminal domain houses the conserved motifs characteristic of all mammalian and even prokaryotic DNA-(cytosine-5)-methyltransferases. This homology with bacterial DNMTs is interesting, particularly in light of the endosymbiotic theory, which proposes that mitochondria are remnants of prokaryotes themselves. The endosymbiotic theory proposes that mitochondria originated as free-living bacteria that were engulfed, and eventually propagated, by eukaryotic cells. This symbiotic union was tolerated because the acquisition of mitochondria by a eukaryotic cell provided some advantage, for example, more efficient energy production, and over time, most of the genes encoded in this symbiotic prokaryote were transferred to the nucleus. As a result, the symbiotic prokaryote or mitochondrion precursor became dependent on the transfer of nuclear proteins and enzymes to assist in complete functioning of the organelle. It is both tempting and reasonable to speculate that the mtDNMT1 isoform is among the proteins that originated in a bacterial symbiont, was transferred to the host nucleus, where it acquired additional sequence and structure including a MTS, and is now targeted back into mitochondria to maintain methylation of mtDNA.

As discussed in Chapter 1, the molecular structure of DNMT1 is thought to have arisen through the fusion of three separate genes: one gene that encoded the catalytic domain and shares strong homology with bacterial methyltransferases, and two genes that together form the regulatory domain (71). It is possible that what is now regarded as the C-terminal catalytic domain was at one time the full active form of DNA methyltransferase within the symbiotic prokaryote. Hypothetically, upon invasion of a host organism, the prokaryotic origin of DNMT1

was amended for a specific purpose within the host nucleus by adding a nuclear localization signal, DNA binding domains and other domains for efficient functioning in the nucleus. Over time, it became apparent that the symbiotic prokaryote also required a functional DNA methyltransferase, thus a mitochondrial targeting sequence was added to allow proper translocation of the modified precursor DNMT1 back into what is now a subcellular organelle: the mitochondrion.

In keeping with this idea, bacterial DNA is thought to be methylated as a protective measure, to prevent its restriction by endonucleases. If mitochondria are truly remnants of prokaryotes, then perhaps methylation of mtDNA has evolved along with the organelle, as a means of protecting it from intra-mitochondrial endonuclease cleavage.

Chapter 3: Generation of a TAP-tagged DNA methyltransferase 1 (DNMT1) cell line

INTRODUCTION

I. Objectives

Central to understanding the function of DNMT1 in the mitochondrial compartment, was the need for a means of efficiently isolating the enzyme under a variety of experimental conditions, including its native context. We needed a straightforward way to ask whether mtDNMT1 binds and methylates mtDNA, and which accessory proteins are required for its function in the mitochondria.

II. The Tandem Affinity Purification (TAP) tag

An efficient method for isolating native protein complexes was initially developed in yeast (107, 108) and since, has been applied successfully to a number of other organisms, including mammalian cells. The method is called tandem affinity purification (TAP), and involves fusion of the TAP-tag in-frame with a gene of interest, allowing expression of a hybrid fusion protein. The TAP-tag contains an IgG-binding domain of the *Staphylococcus aureus* protein A (ProtA) and a calmodulin-binding peptide (CBP) domain, separated by a tobacco etch viral (TEV) protease cleavage site (Figure 3-1A). The combination of ProtA and CBP tags allows for two high-affinity purification steps, to achieve greater purity than either step alone (109, 110). Additionally, the entire procedure is carried out under mild, non-denaturing conditions so that the purified protein should retain functionality. This was critical for our

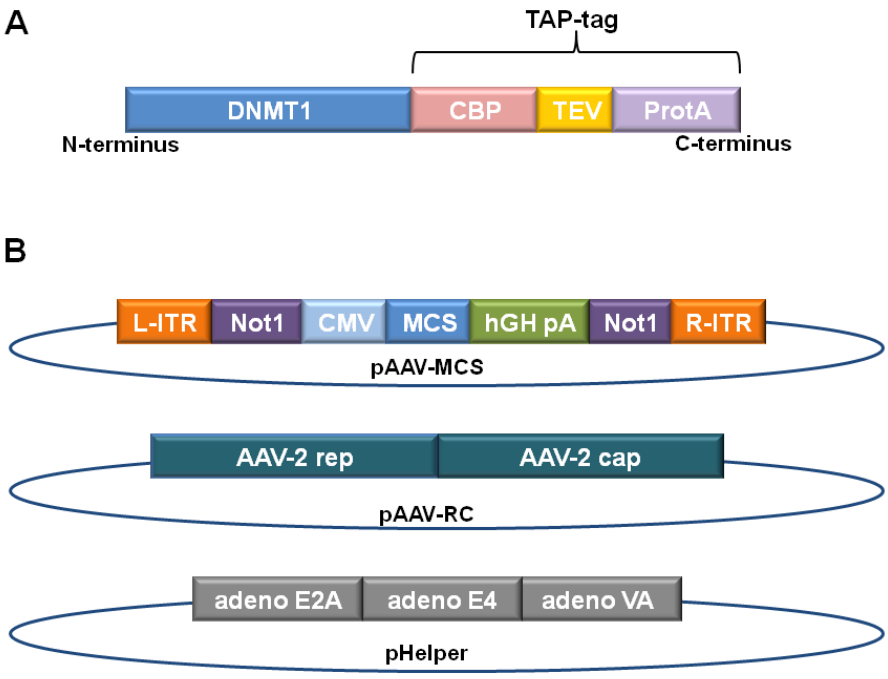


Figure 3-1: Molecular structure of the TAP-tag and rAAV Helper Free system used for gene targeting. (A) The tandem affinity purification (TAP)-tag consists of a Protein A (ProtA) domain and a calmodulin-binding peptide (CBP) domain, separated by a tobacco etch viral (TEV) protease cleavage site. The TAP-tag was fused to the C-terminus of DNMT1, just upstream of the endogenous stop codon for DNMT1. (B) The rAAV Helper Free system (Stratagene) consists of the pAAV-MCS plasmid for direct insertion of your target gene into the multiple cloning site (MCS). The pAAV-MCS vector is not infectious alone, however, when co-transfected with the pAAV-RC and pHelper plasmids, it acquires the genes and accessory proteins necessary for efficient packaging and replication of the recombinant adeno-associated virus.

analyses, because we hoped to determine whether the mtDNMT1 isoform possessed catalytic activity and might be responsible for methylating mtDNA.

The TAP-tag purification protocol is depicted in Figure 3-2. Briefly, the ProtA domain of the tagged protein binds tightly to IgG sepharose beads. Since ProtA bound by matrix IgG can only be released at low pH under denaturing conditions, a specific tobacco etch viral (TEV) protease cleavage site was conveniently placed between the two affinity tags (109, 110). Incubation with TEV protease cleaves the tag, allowing elution of the target protein under native conditions. The eluate of this first affinity purification is then incubated with calmodulin sepharose beads in the presence of calcium to allow a tight calmodulin-CBP interaction. A series of washes removes contaminants, including TEV protease, and the bound material is released through the addition of EGTA, which chelates the calcium required for the calmodulin-CBP interaction.

Prior to development of the TAP-tag and its associated purification method, comprehensive studies examining protein assemblies were limited to *in vitro* and *ex vivo* systems, which often rely on expression of a target protein at levels far from physiologically relevant. TAP-tag purification paired with mass spectrometry (TAP-MS) is an extremely powerful method for determining the identity of unknown proteins in complex with a purified protein of interest, all expressed at natural levels. Several large-scale studies employing the TAP-tag methodology provided insight into how the yeast proteome is organized (107, 111, 112). Proteins of interest were endogenously tagged at their 3' ends with the TAP-tag, purified from total cellular lysates by the TAP-tag method, and resolved on denaturing SDS-PAGE gels. Protein bands were digested with trypsin and analyzed by matrix-assisted laser

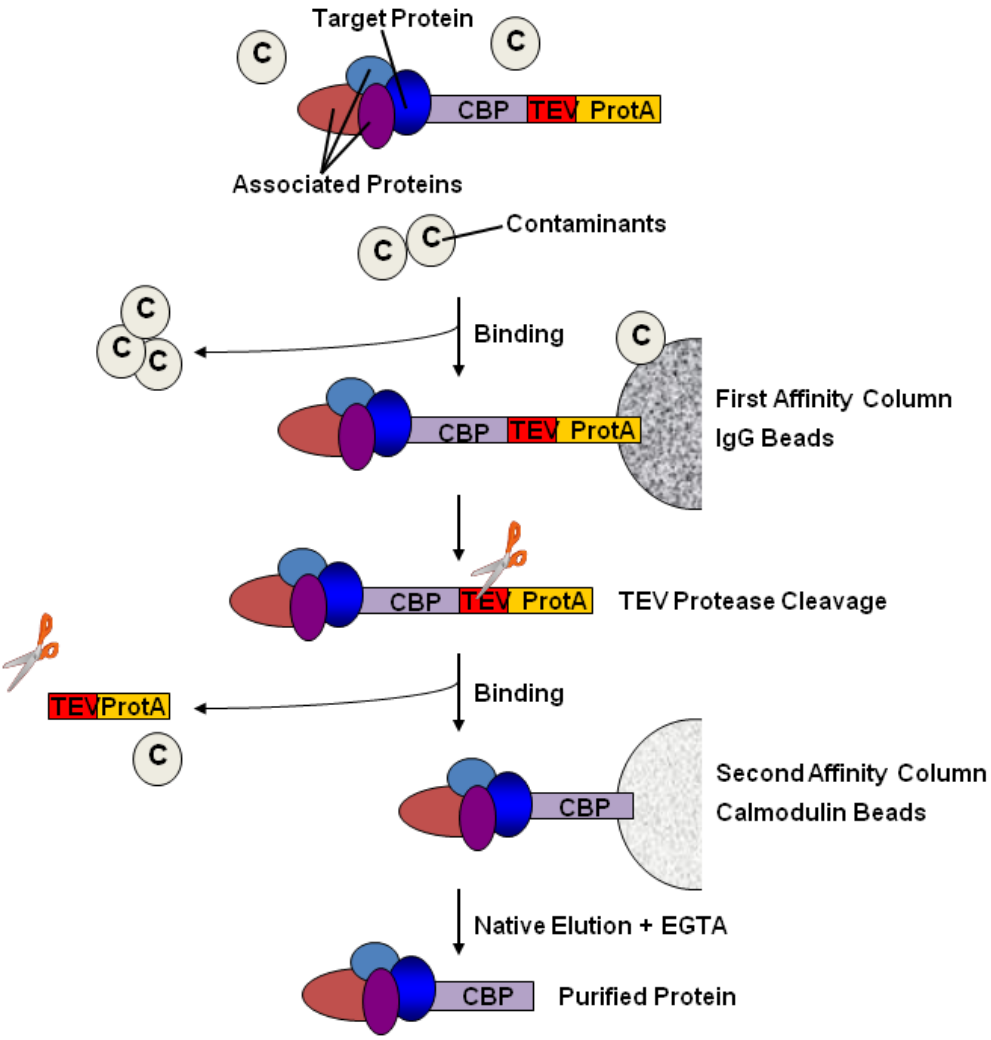


Figure 3-2: Schematic diagram of TAP-tag purification. Purification of a TAP-tagged protein is achieved in four simple steps: 1) Binding of the tagged protein to IgG sepharose, 2) Release of the bound protein by TEV protease cleavage, 3) Binding of the now CBP-tagged protein to calmodulin-coated beads in the presence of high [Ca2+], and 4) Elution of the bound protein with a buffer containing a divalent cation chelating agent, such as EGTA.

desorption/ionization-time-of-flight mass spectrometry (MALDI-TOF MS); their identities were predicted by database searches and algorithms (107).

These experiments generated libraries of over 1,500 yeast strains, and purified nearly 2,000 TAP-fusion proteins, which are estimated to represent approximately 60% of the entire yeast proteome, spanning all functional classes and subcellular localizations of proteins (107, 111). Of the TAP-fusion proteins analyzed, the majority (78-88%) were purified in association with at least one protein binding partner, demonstrating the utility and efficiency of this approach for purifying protein complexes (107, 111). Protein complexes were observed ranging in size from 2-83 components, which illustrates the ability of the TAP approach to purify even large, multi-subunit protein complexes (107).

The TAP-MS method was shown to be highly reproducible, with a 70% probability of detecting the same protein in duplicate preparations (107, 111). Sensitivity of the TAP-MS approach is also high, with limits of detection and identification of proteins present as low as 15 copies per cell (107). In only a small percentage (18%) of attempts did TAP-tagging the C-terminus of a protein generate a non-viable yeast strain, suggesting that insertion of the TAP-tag impaired protein function (107).

From the data collected, researchers were able to assign functional roles to whole complexes, based on the known functions of its individual protein components. There was a wide distribution of predicted cellular roles for the purified complexes, including roles in the cell cycle, transcription/DNA maintenance/chromatin structure, protein/RNA transport, and energy metabolism, among others (107).

Perhaps the most valuable information obtained from these systematic analyses was in the identification of novel protein-protein interactions. These studies were able to propose

cellular roles for proteins that had no known function, and new roles for known proteins. Together, the data paints a clearer picture of how eukaryotic proteins exist in time and space, and how they are organized into higher-order functional networks. Since the advent of the TAP-MS approach in yeast, numerous other groups have used this technique to endogenously tag various epitopes within human somatic cells, including phosphatase and tensin homolog (PTEN), signal transducer and activator of transcription 3 (STAT3) and p53, among others (113, 114).

III. Gene targeting in somatic cells

Among the many considerations regarding an optimal method for delivery of the TAP-tag targeting vector to HCT116 human colon carcinoma cells was that overexpression of DNMT1-TAP was not preferable, particularly because our goals included looking at activity and interactions of the protein, and overexpression is known to lead to the formation of nonspecific interactions and/or false-positive activity readings (108, 109). Dr. Bert Vogelstein and colleagues (115) described an efficient method for gene targeting in somatic cells through homologous recombination. The protocol was adapted from those used successfully in yeast (107, 108), and employed recombinant adeno-associated viruses (rAAV) to achieve 10-100 times higher frequencies of recombination than those obtained with more conventional plasmid vectors (115). The rAAV technology is now marketed and available for purchase through Stratagene, termed the "AAV Helper-Free System". The kit provides the pAAV-MCS vector (Figure 3-1B), into which the targeted insertion/deletion construct is cloned, as well as the pAAV-RC and pHelper plasmids, which encode the *rep* and *cap* genes and accessory adenovirus proteins required for efficient rAAV packaging and replication (115). The simplicity of this approach made construction of the TAP-tagged DNMT1 targeting construct a straightforward exercise in PCR

and cloning, and resulted in a valuable tool that will be useful in a variety of downstream applications.

IV. Destabilization of *DNMT1* mRNA through interaction with AUF1

The C-terminus of DNMT1 was selected for targeted insertion because we wanted the DNMT1-TAP fusion protein to be expressed off of the native DNMT1 promoter, and under the same regulation as the non-tagged endogenous DNMT1 allele. Additionally, the mitochondrial targeting sequence (MTS) for DNMT1 is located within the 5' flanking sequence of the gene (far N-terminus), and remains to be fully characterized, making the C-terminus a more desirable option for tagging. However, among the potential pitfalls of targeting the C-terminus of DNMT1 is that it may interfere with normal regulation of the DNMT1-TAP fusion protein, as DNMT1 has been shown to contain a conserved AU-rich element in the 3'UTR, which mediates stability of the *DNMT1* mRNA (92, 116, 117).

Many human genes, including *c-myc*, *c-fos*, and several cytokines contain AU-rich elements (AREs) that serve as regulatory regions, modulating their mRNA levels post-transcriptionally in response to changing cellular conditions (116, 117). Interaction of these regions with ARE binding factors, such as AU-rich element/poly(U)-binding/degradation factor 1 (AUF1) has been shown to accelerate mRNA decay (117). Szyf and colleagues (92, 116) identified a similar conserved ARE within the 3'UTR of *DNMT1,* which came to our attention as we designed the C-terminal DNMT1-TAP targeting construct. Using RNA immunoprecipitation with an AUF1 antibody, followed by RT-qPCR, Torrisani, *et al.* (92) showed that when this region was bound by AUF1, the *DNMT1* mRNA underwent rapid decay. Overexpression of AUF1 resulted in a decrease in the *DNMT1* mRNA level; alternatively, knockdown of AUF1 led

to increased DNMT1 protein expression (92). The exosome complex was found to be involved in this process, so they concluded that the mechanism underlying AUF1-mediated regulation of DNMT1 is destabilization of the *DNMT1* mRNA by targeting it to the exosome (92).

MATERIALS AND METHODS

Materials

HCT116 neo^S parental cells and HEK293 cells were obtained from American Type Culture Collection (ATCC, Manassas, VA). DNA oligo primers (25nmoles) were designed and purchased from Integrated DNA Technologies (IDT, Coralville, IA). All tissue culture was performed in a sterile, laminar-flow tissue culture hood in a designated room, separate from the rest of the laboratory. The pZome-1C vector, containing the tandem affinity purification (TAP) tag was obtained from Dr. Bertrand Seraphin, European Molecular Biology Lab. The pNeDaKO-Neo vector, housing the Neo^R cassette used for selection of successfully-transfected clones was obtained from Dr. Bert Vogelstein, Johns Hopkins University. The pMC1-Cre plasmid was obtained from Dr. Jolene Windle, VCU, and purified by Cortney Heyer in Dr. Moran's laboratory. 2x iProof GC-rich Mastermix (BioRad, cat # 172-5320) was used for all endpoint PCR, unless otherwise stated. A BioRad Peltier Thermal Cycler (DNA Engine, model PTC-200) was used for all endpoint PCR reactions with cycling conditions described for each primer set separately. A 10µL aliquot of the 1kb+ DNA ladder (Invitrogen, cat # 10787-018) was used to determine the size of all electrophoresed DNA products. The Promega Wizard SV gel and PCR clean-up kit (cat # A9282) was used to purify all PCR products excised from agarose gels; the "DNA Purification by Centrifugation" protocol was followed as described in the manufacturer's instructions. Quantitect SYBR Green Mastermix (cat# 204145) was purchased from Qiagen and

used for all qPCR analysis. BioRad DNA Engine Peltier thermal cyclers, fitted with Chromo 4 Real-Time Fluorescence Detector attachments were used for all studies involving qPCR. Real-time qPCR data was analyzed with Opticon Monitor 3 Software. All restriction endonucleases were purchased from New England Biolabs (Ipswich, MA) unless otherwise stated. Antibodies were purchased from Abcam (Cambridge, MA), Sigma (St. Louis, MO), Upstate Biotechnology (Billerica, MA) and Open Biosystems (Rockford, IL), and will be discussed individually below. Radioactivity in Southern blots and DNMT1 enzymatic assays was measured using a Beckman Coulter LS 6500 Multi-purpose liquid scintillation counter with programs specific to each isotope.

Methods

Cell Culture

HCT116neo^S cells and HCT116-TAP cells were grown at 37°C in 5% CO_2 inRPMI 1640 (Gibco/Invitrogen) medium, supplemented with 10% FBS, 100U/mL Penicillin and 100mg/mL Streptomycin. HEK293 cells were grown at 37°C in 10% CO_2 in DMEM (Gibco/Invitrogen) medium, supplemented with 10% FBS, 100U/mL Penicillin and 100mg/mL Streptomycin. Cell cultures were fed every other day, and stock cultures were maintained at subconfluent cell densities.

For experimental consistency and because DNMT1 is cell-cycle regulated (78, 92), all cells were treated the same way in the days leading up to an experiment. Cells were fed 48 hrs prior to harvest, trypsinized and seeded 24 hrs prior to harvest to ensure that the majority of cells were still in S-phase, and expressing the maximal level of DNMT1.

Generation of the DNMT1-TAP-tag targeting construct

Amplification and gel purification of homology arms

Approximately 950bp of sequence homologous to coding exons 39 and 40 and intron 39 from human *DNMT1* comprised the left homology arm (LHA), and approximately 900bp of sequence homologous to the 3' UTR (Exon 40) and the 3' flanking sequence from the human *DNMT1* gene comprised the right homology arm (RHA). These stretches of 900-950bp matching the endogenous C-terminus of DNMT1 provide sufficient sequence (113) to allow for homologous recombination of the targeting vector with the genomic DNMT1 locus, thereby inserting the TAP-tagged DNMT1 sequence into the genome.

Both the LHA and RHA were amplified against 20ng genomic DNA from HCT116 cells using primers P1sn and P2asn for LHA or primers P3sn and P4asn for RHA (Table 3-1) in a 25µL reaction containing 12.5µL of 2x iProof GC-rich Mastermix, 0.6µM of each primer, and 3% DMSO. The PCR program consisted of initial denaturation at 94°C for 30sec, and 35 cycles each including: denaturation at 94°C for 30sec, annealing at primer set-specific temperatures for 30sec, extension at 72°C for 1min, followed by final extension at 72°C for 5min and a hold at 4°C.

Both P1sn and P4asn primers were designed to include *Not*I restriction sites (Figure 3-3), such that these sites would be located at the far 5' and 3' ends of the final targeting construct, for digestion and ligation into the pAAV-MCS vector (Figure 3-1). The P2asn primer contained 24bp of sequence complementary to the TAP-tag, and 21bp of sequence complementary to the 3' end of the LHA, excluding the endogenous "stop" codon for DNMT1. The P3sn primer contained 23bp of sequence at the 3' end of the Neo^R cassette (labeled as Linker "B" in Figure 3-

Table 3-1: TAP-tagging DNMT1 primers

TAP-tagging DNMT1 Primers		
P1sn	5'-ATACATACGCGGCCGCCCCACGTGTCTTTGTCTCAAGTCTTTC-3'	
P2asn	5'-CCATCTTCTCTTTTCCATGGATCCGTCCTTAGCAGCTTCCTCCTC-3'	
P3sn	5'-CGCCCTATAGTGAGTCGTATTACTTCTGCCCTCCCGTCACCCCTGTTT-3'	
P4asn	5'-ATACATACGCGGCCGCAGGAGGCAGAGGCTGCATTGAACG-3'	
TAPsn	5'-GGACGGATCCATGGAAAAGAGAAGATGG-3'	
TAP-LinkerA-asn	5'-GCTCCAGCTTTTGTTCCCTTTAGCTAATTAGCGTCTACTTTCGGCGC-3'	
P5a	5'-AGCGAAACGTGTCTAGGTGGCC-3'	
TAPrev	5'-TAGCGCTTTGGCTTGGGTCATCTT-3'	
P6b	5'-TCCAGTGGCCAAAGGACACACAGAAG-3'	
Sequencing		
LHA-TAPseq	5-CGGGCGTGTTCCCCAGAGTGAC-3'	
Neo sn	5'-TGGGCTCAGAGGCTGGGAAGGGGTG-3'	
Primer Set	**Forward**	**Reverse**
Amplification of Southern Blot Probes		
SB LHA2	5'-CCGTATGCTGTCACAGTGCCATTT-3'	5'-TTAAGCCAACCCGACACCTGAGAA-3'
SB RHA2	5'-GCCTATTTAGAAACGCCTTTCCTCCC-3'	5'-GATCAAGGTGCCTTGGAATCCAGAA-3'
Quantitative PCR		
WT	5'-TGCTAAGGACTAGTTCTGCCC-3'	5'-CCTCACACAACAGCTTCATGTCAGC-3'
TAP	5'-TGCTAAGGACGGATCCATG-3'	5'-TTGTTGTCCACGGCTTCATC-3'
Internal DNMT1	5'-TCCGAGATGCCGGCGCGTACC-3'	5'-CTCTTTCCAAATCTTGAGCCGCC-3'
Human β-actin	5'-CACCACACCTTCTACAATGAG-3'	5'-GGTCTCAAACATGATCTGGGTC-3'
RNA Immunoprecipitation		
hFOS	5'-GAGAGCTGGTAGTTAGTAGCATGTTGA-3'	5'-AATTCCAATAATGAACCCAATAGATTAGTTA-3'
Ex37-38	5'-TTCTTCAGCACAACCGTCACCAAC-3'	5'-TTGTCCAGGATGTTGCCGAAGA-3'

3, which was required for subsequent Fusion PCR reactions) and 25bp of sequence spanning the end of Exon 40 into the 3' UTR for DNMT1.

PCR products for the LHA and RHA amplication reactions were electrophoresed on a 1% TAE-EtBr agarose gel alongside the 1kb+ DNA standard ladder, and visualized by UV transillumination. PCR products of the correct size (~950bp for LHA; ~900bp for RHA) were excised from the gel using a sterile razor blade and purified using the Promega Wizard SV gel and PCR clean-up kit. Purified PCR products were quantitated on the NanoDrop spectrophotometer, and stored at -20°C.

Amplification of the TAP-tag from the pZome-1C vector

The TAP-tag is comprised of sequence encoding a Protein A domain (ProtA), linked in tandem to sequence encoding a calmodulin binding peptide (CBP), and between them lies a cleavage site for tobacco etch viral (TEV) protease (Figure 3-1). The TAP-tag was amplified using 50pg of the pZome-1C plasmid with primers TAPsn and TAP-LinkerA-asn (Table 3-1) in a 25µL reaction containing 12.5µL of 2x iProof GC-rich Mastermix, 0.6µM of each primer, and 3% DMSO. The PCR program consisted of initial denaturation at 94°C for 30sec, and 35 cycles each including: denaturation at 94°C for 30sec, annealing at 50°C for 30sec, extension at 72°C for 30sec, followed by final extension at 72°C for 5min and a hold at 4°C. The TAPsn primer contained 24bp of sequence complementary to the P2asn primer, which was required for subsequent Fusion PCR reactions. The TAP-LinkerA-asn primer contained 23bp of sequence complementary to the 5' end of the Neo^R cassette, and 24bp of sequence complementary to the "Linker A" region, located at the 3' end of the TAP-tag.

PCR products for the TAP-tag amplication reactions were electrophoresed on a 1% TAE-EtBr gel alongside 10µL of 1kb+ DNA standard ladder, and visualized by UV transillumination

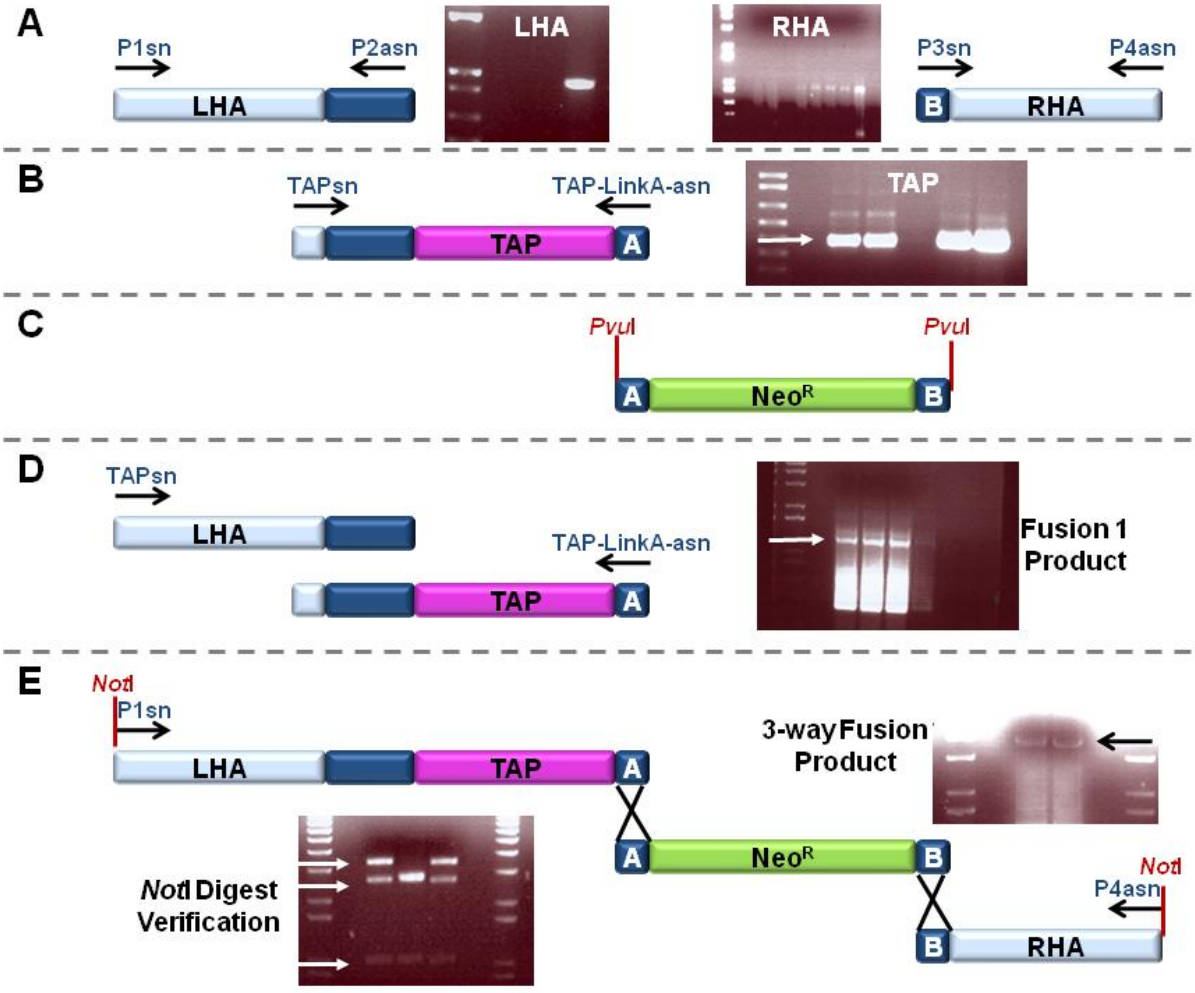


Figure 3-3: Fusion PCR strategy to generate the TAP-tagged DNMT1 targeting construct. (A) Left and right homology arms (LHA and RHA, respectively) were amplified from the genomic locus of DNMT1 using total gDNA from HCT116 cells as a template. Primers are denoted by arrows; primer sequences can be found in Table 3-1. (b) The TAP-tag was amplified using the pZome-1C plasmid as a template. (C) The NeoR selection cassette was excised from the pNeDaKO-Neo vector by restriction digest with the *Pvu*I endonuclease. (D) The LHA was fused with the TAP-tag by PCR using the TAPsn primer, paired with the TAP-LinkA-asn primer to generate a ~1500bp fragment representing the Fusion 1 product. (E) The final DNMT1-TAP targeting construct was created by 3-way fusion PCR, generating a 4.5kb fragment for insertion into the pAAV-MCS vector. A *Not*I restriction digest was performed to verify whether clones contained the full insert in the correct orientation.

(Figure 3-3). PCR products of the correct size (~550) were purified, quantitated on the NanoDrop spectrophotometer, and digested with *Sac*I restriction enzyme (NEB, cat # R0156S) to verify the identity of the PCR product. *Sac*I digestion was performed in a 25µL reaction, containing 250ng purified TAP-tag DNA template, 20U (1µL) *Sac*I enzyme, and 2µL 10x NEB Buffer 1. Digest products were electophoresed on a 1% TAE-EtBr agarose gel alongside the 1kb+ DNA ladder and visualized by UV transillumination. Two bands of the expected sizes were observed, indicating that we had purified the correct TAP-tag PCR product.

Restriction digest to excise the Neo^R cassette from the pNeDaKO-Neo vector

The pNeDaKO-Neo vector was digested with *Pvu*I restriction endonuclease (NEB, cat # R0150S) in a 20µL reaction containing the following: 6.5µg pNeDaKO-Neo plasmid, 10U (1µL) *Pvu*I enzyme, 2µL 10x NEB Buffer 3, 0.2µL 10x BSA. The reaction was incubated at 37°C for 1hr. Digestion products were electrophoresed on a 1% TAE-EtBr agarose gel alongside the 1kb+ DNA standard ladder, and visualized by UV transillumination. The expected 3kb band was purified.

Fusion1 PCR to link the LHA with the TAP-tag

Equimolar amounts of the purified LHA and TAP-tag templates were combined in a 25µL reaction containing 12.5µL 2x iProof GC-rich Mastermix, 0.6µM P1sn primer, and 0.6µM TAP-LinkA-asn primer. The PCR cycling program consisted of initial denaturation at 94°C for 30sec, followed by 35 cycles of the following: denaturation at 94°C for 30sec, annealing at 58°C for 30sec, and extension at 72°C for 1min, followed by final extension at 72°C for 5min, and a hold at 4°C. PCR products were electrophoresed on a 1% TAE-EtBr agarose gel alongside the 1kb+ DNA standard ladder , and visualized by UV transillumination. Faint bands of the expected size of 1500bp (950bp LHA + 550bp TAP) were observed in both lanes (Figure 3-3D).

Adenylation of the LHA-TAP Fusion 1 PCR product was achieved by combining a maximal volume of the PCR product with 1U (0.2µL) Taq Polymerase and 1µL dATP in a 25µL reaction, and incubating at 72°C for 15min. Adenylated products were electrophoresed on a 1% TAE-EtBr agarose gel alongside the 1kb+ DNA standard ladder and visualized by UV transillumination. The LHA-TAP Fusion PCR product + A's migrated at the expected size of 1500bp, thus bands of the appropriate size were excised from the gel and purified using the Promega Wizard SV gel and PCR clean-up kit.

3-way Fusion PCR to create the final targeting construct

A total of 350ng of the purified LHA-TAP Fusion 1 PCR product, 350ng of the purified RHA PCR product, and 400ng of the purified Neo^R cassette, representing the three DNA templates to be fused in this PCR reaction (Figure 3-3E), were combined into a 400µL mastermix, containing the following: 200µL of 2x iProof GC-rich Mastermix, 0.6µM P1sn primer, and 0.6µM P4asn. The large 400µL reaction was mixed thoroughly by vortexing and quickspun to collect all contents at the bottom of the tube, before aliquoting 50µL into the eight individual 0.2mL wells of a TempAssure 8-tube PCR strip (USA Scientific, cat # 1402-2700). The PCR cycling program consisted of initial denaturation at 94°C for 2min, followed by 20 cycles, each including: denaturation at 94°C for 30sec, annealing at 56°C for 30sec, extension at 68°C for 4min, followed by final extension at 68°C for 5min and a hold at 4°C.

*Not*I digestion of 3-way Fusion PCR product and pAAV-MCS vector

The eight PCR tubes containing 3-way fusion products were pooled into four 1.5mL microcentrifuge tubes (each containing 100µL) for enzymatic digestion with *Not*I restriction endonuclease. For each of four reactions, the following was combined in a 200µL reaction: 100µL 3-way Fusion PCR products, 60U (1.2µL) *Not*I restriction enzyme, 20µL 10x NEB

Buffer 3, and 8µL 10x BSA. The reactions were pulse-vortexed to mix and quick-spun to ensure all contents were at the bottom of the tubes, before incubating at 37°C for 3hrs. All four digested samples were pooled into two 1.5mL microcentrifuge tubes for Phenol/Chloroform extraction and EtOH precipitation, performed as described in Chapter 2 for MeDIP. Each DNA pellet was resuspended in 20µL of TE, and mixed thoroughly to ensure the DNA pellet was fully solubilized. All 40µL of precipitated DNA was loaded onto a 0.8% TAE-EtBr agarose gel, electrophoresed and visualized by UV transillumination. Bands of the expected size (~4.5kb) appeared in both sample lanes (Figure 3-3), and were excised and purified. Purified DNA products were quantitated on a NanoDrop spectrophotometer, and stored at -20°C.

To linearize the pAAV-MCS plasmid and prepare it for insertion of the TAP-tagged DNMT1 construct, 2µg of the circular pAAV-MCS plasmid was combined with 75U (1.5µL) *Not*I restriction endonuclease (NEB, cat # R0189M), 2µL 10x NEB Buffer 3, and 0.2µL BSA in a 20µL reaction. The reaction was incubated at 37°C for 2hrs, before 20U (2µL) of calf intestinal phosphatase was added, and incubated for an additional 15min to dephosphorylate the vector. Restriction digest products were electrophoresed on a 1% TAE-EtBr agarose gel alongside the 1kb+ DNA ladder, and visualized by UV transillumination. A band of the expected 3kb size was observed, excised from the gel, and purified using the Promega Wizard SV kit according to manufacturer's instructions.

Cloning of the DNMT1-TAP targeting construct into pAAV-MCS

To ligate the purified, *Not*I-digested 3-way Fusion PCR product into the *Not*I-digested pAAV-MCS vector for cloning, the following was combined in a 20µL ligation reaction: 70ng (12.5µL) of *Not*I-digested 3-way Fusion PCR product, 11ng (0.5µL) of the linearized pAAV-MCS plasmid, 2µL of 10x T4 DNA Ligase Buffer and 1µL of T4 DNA Ligase (Invitrogen, cat #

15224-017). A no-insert negative control ligation reaction was also set up using 12.5µL of H2O instead of insert. Both reactions were pulse-vortexed to mix, and quickspun to collect all contents at the bottom of the tube. Ligation reactions were incubated at room temperature for 30min, then placed on ice for 5min. Half of each ligation reaction (10µL) was used in a transformation reaction, the other half (10µL) was stored at 4°C as a back-up.

Transformation of the ligated pAAV-MCS-DNMT1-TAP vector was achieved using OneShot TOP10 Chemically-competent *E.coli* cells (Invitrogen, cat # C404010). Two vials of TOP10 cells were removed from storage at -80°C and thawed on ice. SOC Medium, supplied with the TOPO TA cloning kit (Invitrogen, cat # K4500-01) was removed from storage at -20°C and warmed in a water bath set to 42°C until use. The pUC19 positive control plasmid for transformation was also removed from storage at -20°C and thawed on ice. 10µL of the pAAV-MCS-DNMT1-TAP ligation reaction was added to 1 vial (~50µL) of TOP10 cells, and the tube was tapped gently to mix its contents. To half of the other vial of TOP10 cells, 10µL of the no-insert negative control ligation reaction was added, and the tube was tapped gently to mix. Lastly, to the remaining half vial of TOP10 cells, 1µL of the pUC19 positive control vector was added, and the tube was tapped gently to mix. All transformation reactions were incubated on ice for 30min, then heat-shocked for 45sec in a 42°C water bath. The transformations were placed on ice for 2-5min, and then 250µL of pre-warmed SOC Medium was added directly to each vial. The vials were placed in a heated horizontal shaker, shaking at 225rpm at 37°C for 1hr.

To prepare LB agarose plates for plating the transformation reactions, LB agarose (10g Bacto-tryptone, 5g yeast extract, 10g NaCl, and 15g agar in 1L ddH2O) was melted and allowed to cool almost to room temperature. To 48mL of LB agar, 40µg/mL of Zeomycin and 60µg/mL of Carbenicillin was added and swirled to mix. Approximately 12mL of LB agar plus Zeomycin

(Zeo) and Carbenicillin (Carb) selection agents was poured into each bacterial petri dish and allowed to solidify. 125µL of each tranformation reaction was added dropwise to each Carb-Zeo plate, and spread with a sterile glass rod until fully absorbed by the agar. Inoculated plates were incubated at 37°C in bacterial incubator overnight. 10 colonies were picked from each ligation plate (20 colonies total) for further analysis. Liquid cultures of each colony were made by inoculating 5mL of LB + Carb + Zeo liquid broth. Cultures were placed in 14mL round bottom conical tubes (BD Falcon, cat # 352059) with the caps loose to allow sufficient oxygen exchange, and incubated in a heated horizontal shaker at 37°C, shaking at 225rpm overnight.

Purification and verification of DNA from pAAV-MCS-DNMT1-TAP clones

DNA was purified from liquid cultures of pAAV-MCS-DNMT1-TAP transformed colonies using the Promega Wizard Plus SV Miniprep DNA Purification System (cat # A1460) following the centrifugation protocol according to manufacturer's instructions. Two elution steps, 50µL each, were used to collect the purified DNA from the spin columns. Samples were quantitated on the NanoDrop spectrophotometer, and stored at -20°C until further analysis.

DNA samples were thawed, pulse-vortexed to mix, and quickspun to collect all contents at the bottom of each tube. A restriction digestion with *Not*I endonuclease was set up to verify the identity of each DNA sample. For each 20µL digest, the following was combined: 250ng of Miniprepped DNA product, 15U (0.3µL) of *Not*I enzyme, 2µL of 10x NEB Buffer 3, and 0.2µL of 10x BSA. Tubes were tapped and quickspun to be sure that all reaction contents were well-mixed and at the bottom of each tube, and reactions were incubated at 37°C for ~2.5hrs. Digested products were electrophoresed on a 0.8% TAE-EtBr agarose gel and visualized by UV transillumination (Figure 3-3E). Two of the 20 colonies were shown to contain the expected restriction digest products: a 4.5kb band, representing the excised DNMT1-TAP insert, and a 3kb

band, which represents the linearized pAAV-MCS vector. A third colony was questionable, as it produced bands of 3kb and 4kb; the latter product was slightly smaller than expected for the full insert. A second restriction digest was performed to ensure that the clones contained the right sequence. *Fsp*I restriction endonuclease (NEB, cat # R0135S) was chosen for this second digest. *Fsp*I cut three times within the pAAV-MCS-DNMT1-TAP targeting vector sequence: once within the vector, once within the TAP-tag sequence, and once within the Neo^R cassette, generating three separate fragments. Minprepped DNA (250ng) from the three colonies of interest were added to 20µL digestion reactions, combined with 1.5U (0.3µL) *Fsp*I enzyme, and 2µL of 10x NEB Buffer 4. Tubes were tapped and quickspun to be sure that all reaction contents were well-mixed and at the bottom of each tube, and reactions were incubated at 37°C for ~1.5hrs. Digested products were electrophoresed on a 0.8% TAE-EtBr agarose gel and visualized by UV transillumination. Two of the three clones (1b and 2c) produced bands of the expected sizes: 1.1kb, 2.5kb and 3.5kb, indicating that they did indeed contain the correct targeting construct (Figure 3-3). DNA from these two samples was prepared for sequencing by the VCU Nucleic Acids Research Facility, according to their specifications.

Construction of a positive control plasmid for screening recombinant DNMT1-TAP clones by PCR

In order to screen the resultant HCT116 TAP-tagged DNMT1 clones for proper integration of the targeting construct, we first needed to create a positive control plasmid containing the endogenous DNMT1 sequence extending *outside* of the insertion site, fused to one of the artificial knock-in elements (i.e. the TAP-tag). This would allow us to PCR amplify and sequence across the sites of integration of the DNMT1-TAP targeting construct to directionally

verify its proper insertion into the genomic locus of DNMT1 in each of our G418-resistant clones.

The genomic sequence for DNMT1 surrounding the insertion site of the DNMT1-TAP targeting construct, from Exon 38 through the 3'UTR, was analyzed by Repeat Masker (www.repeatmasker.org) to highlight stretches of sequence that were unique to human DNMT1, so that stretches of highly repetitive sequence could be avoided during primer design. Using this information, new primers were designed to fall only within the regions specific to DNMT1. The P5a primer (Table 3-1) was designed against sequence located >100bp upstream of the LHA. A series of enrichment PCR reactions were run such that the product purified from one PCR reaction was used as the template in the next PCR reaction, all using the same set of primers (P5a and P2asn). An aliquot (100µL) of an enrichment PCR reaction, containing: 50µL of 2x iProof GC-rich mastermix, 0.6µM primer P5a, 0.6µM primer P2asn, and 30ng (10µL) purified Extended LHA PCR product, was assembled, mixed thoroughly and divided into four 25µL reactions. The PCR cycling program consisted of initial denaturation at 94°C for 2min, followed by 35 cycles, each including: denaturation at 94°C for 30sec, annealing at 56-59°C for 30sec, extension at 72°C for 1min, followed by final extension at 72°C for 5min and a hold at 4°C. PCR products were resolved by gel electrophoresis on a 1% TAE-EtBr agarose gel, and visualized by UV transillumination. All four sample lanes showed bright bands at the expected 1.2kb size, representing the ExtLHA positive control products. These bands were excised from the gel using a clean razor blade, purified using the QIAquick Gel Extraction kit (Qiagen, cat # 28706) and quantitated on the NanoDrop spectrophotometer.

This new ExtLHA product next needed to be fused with the TAP-tag, exactly as before for the LHA-TAP Fusion 1 PCR reaction. Equimolar amounts of the ExtLHA template and the

purified TAP-tag template, were combined in a 100µL PCR reaction, containing 50µL of 2x iProof GC-rich mastermix, 0.6µM P5a primer, and 0.6µM TAP-LinkA-asn primer. These reagents were mixed well and aliquoted into four separate PCR tubes. The PCR cycling conditions consisted of initial denaturation at 94°C for 2min, followed by 35 cycles, each including: denaturation at 94°C for 30sec, annealing at 56-59°C for 30sec, extension at 72°C for 1min, followed by final extension at 72°C for 5min and a hold at 4°C. PCR products were resolved by electrophoresis on a 1% TAE-EtBr agarose gel and visualized by UV transillumination. All four samples lanes showed bright bands of the expected 1.75kb size. These fragments were excised from the gel and purified using the QIAquick Gel Extraction kit (Qiagen).

Because the iProof GC-rich enzyme (BioRad) generates blunt-ended PCR products, the purified 1.75kb ExtLHA-TAP fragment was adenylated at both ends by combining the following into a 25µL reaction: 20.55µL purified ExtLHA-TAP PCR product, 1U (0.2µL) Taq polymerase (Invitrogen, cat # 10342-053), 200nM dATP, 5µL of 10x PCR Buffer, and 1.5mM $MgCl_2$. The reaction was incubated at 72°C for 15min, then put into a ligation reaction for cloning using the TOPO TA Cloning Kit for Sequencing (Invitrogen, cat # K4575-01), according to the manufacturer's instructions. Ligation, transformation and cloning of the positive control ExtLHA-TAP fusion insert was performed as described previously for the pAAV-MCS-DNMT1-TAP targeting vector. DNA was purified from 10 colonies surviving Carb selection and verified by restriction digest using *Pst*I enzyme (NEB, cat # R0140S), in a 20µL digestion reaction containing: 2.0µL of 10x NEB Buffer 3, 2.0µ 10X BSA, ~300ng (2.0µL) purified ExtLHA-TAP positive control plasmid DNA, 20U (1.0µL) *Pst*I enzyme. The reaction was pulse-vortexed to mix well and quick-spun to collect all contents at the bottom of the tube before

incubating at 37°C for 1hr. Digested products were electrophoresed on a 1% TAE-EtBr agarose gel and visualized by UV transillumination. For all 10 colonies, bands were observed at sizes suggesting that the insert is backwards in its orientation. While this does not matter for screening purposes, because the primers used for PCR screening will recognize the insert regardless of its orientation, this detail will need to be kept in mind for future experiments, or when considering other possible uses for this positive control plasmid.

Packaging of the pAAV-MCS-DNMT1-TAP viral vector in HEK293 cells

Because the pAAV-MCS vector containing the DNMT1-TAP targeting construct lacks the *cis*-acting elements required for replication and packaging of the adeno-associated virus, co-transfection with the pAAV-RC and pHelper plasmids (Strategene AAV Helper-Free System) provided these elements in *trans* (Figure 3-1B). The Lipofectamine 2000 reagent (Invitrogen, cat # 11668-027) was used to supply the AAV plasmids to human embryonic kidney (HEK293) cells (ATCC) for packaging of the virus. Two separate transfection reactions were prepared under a sterile, laminar-flow tissue culture hood: one each for clone 1b and 2c, which contain verified pAAV-MCS-DNMT1-TAP constructs. In each transfection reaction, 2.5µg of purified DNA from one of the verified pAAV-MCS clones was combined with 2.5µg of pAAV-RC and 2.5µg of pHelper plasmids. These DNA mixtures were dissolved in 1x Opti-MEM (Invitrogen/Gibco, cat # 31985-070) to a final volume of 750µL. In two separate tubes, 54µL of Lipofectamine 2000 was dissolved into 696µL of 1x Opti-MEM, to a total volume of 750µL, and allowed to incubate at room temperature under the hood for 5min. After the 5min incubation, the diluted DNA was combined with the diluted Lipofectamine in a sterile 1.5mL microcentrifuge tube, and incubated at room temperature under the hood for 15min. This incubation period is critical for the

formation of cationic lipid complexes with the viral DNA, for more efficient delivery and expression in the packaging cells.

Two T-75 flasks of HEK293 cells, grown to 50-60% confluence, were prepared for transfection by aspirating-off all old media and washing once with 10mL Hank's Balanced Salt Solution (HBSS) (Invitrogen/Gibco, cat # 14170-112). 10mL of fresh media was replaced, and the DNA-Lipofectamine mixtures were added to the cell medium dropwise, swirling the flask to mix. The cells were incubated at 37°C with 10% CO_2 for 3-4hrs. After the incubation period, the media-DNA-Lipofectamine solution was aspirated-off by vacuum, and 10mL of fresh Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% FBS and 1% Penicillin/Streptomycin (PenStrep) was added. Transfected HEK293 cells were incubated at 37°C with 10% CO_2 for 48hrs to allow packaging of the virus.

After 48hrs, the virus was harvested. Media was aspirated-off, and cells were rinsed once with 1x Phosphate Buffered Saline (PBS), pH 7.4 (Invitrogen/Gibco, cat # 10010-049). The AAV-transfected HEK293 cells were scraped to detach them from the bottom of the tissue culture flask, using a sterile cell scraper (Greiner BioOne, cat # 541070), dipped first into 1x PBS, pH 7.4. The scraped cells were collected and transferred into a 14mL round-bottom conical tube (BD Falcon, cat # 35029) already containing 1mL of 1x PBS, pH 7.4. The flask was rinsed once with 1mL of 1x PBS, pH 7.4, and this wash was collected and transferred to the same conical tube. Three cycles of freeze-thaw were performed to lyse the cells: 10min in a dry ice-70% ethanol bath, followed by 10min in a 37°C water bath. Tubes were tapped to mix their contents after each thaw cycle. After lysis, the lysate was transferred to 2mL microcentrifuge tubes for clarifying. Lysates were clarified by spinning at 12,000rpm in a tabletop microfuge for

10min at 4°C. The supernatant was removed and aliquoted into five pre-labeled 1.5mL microcentrifuge tubes (330µL each), and aliquots were stored at -80°C.

Gene targeting in HCT116 parental cells

Generation of the HCT116-TAP cell line was achieved as depicted in Figure 3-4. Two T-75 flasks containing G418-sensitive wild-type HCT116 cells, grown to 60-80% confluence in RPMI 1640 supplemented with 10% FBS, 100U/mL Penicillin and 100mg/mL Streptomycin, were used for gene targeting. Medium was aspirated and cells were washed once with 10mL HBSS, and 4mL of pre-warmed growth medium was added to the flask. One aliquot (330µL) of thawed viral lysate from either verified clone (1b or 2c) was added dropwise to the cell medium in each of two T-75 flasks, swirling the flasks to mix. The virus was allowed to infect cells for 2-3hrs at 37°C, 5% CO_2. After transduction, 8mL of pre-warmed growth medium was added to each flask, and cells were incubated at 37°C, 5% CO_2, for 48hrs before selection.

Selection, screening and verification of original HCT116-TAP cell lines

Plating and selection of recombinants by G418 resistance

Virus-infected HCT116 cells were harvested from both T-75 flasks by trypsinization, and the cell pellets were resuspended in 5mL of fresh growth media. 202mL of selection medium consisting of RPMI 1640, supplemented with 10% FBS, 100U/mL Penicillin, 100mg/mL Streptomycin and a final concentration of 400µg/mL of G418/Geneticin (Gibco/Invitrogen, cat # 11811-031) was prepared and 5mL of resuspended cells was added. Using a multi-channel pipettor fitted with a 300µL head, approximately 200µL of targeted cells in selection medium were seeded into ten 96-well plates (BD Falcon, cat # 353916). Remaining cells (~5mL) were

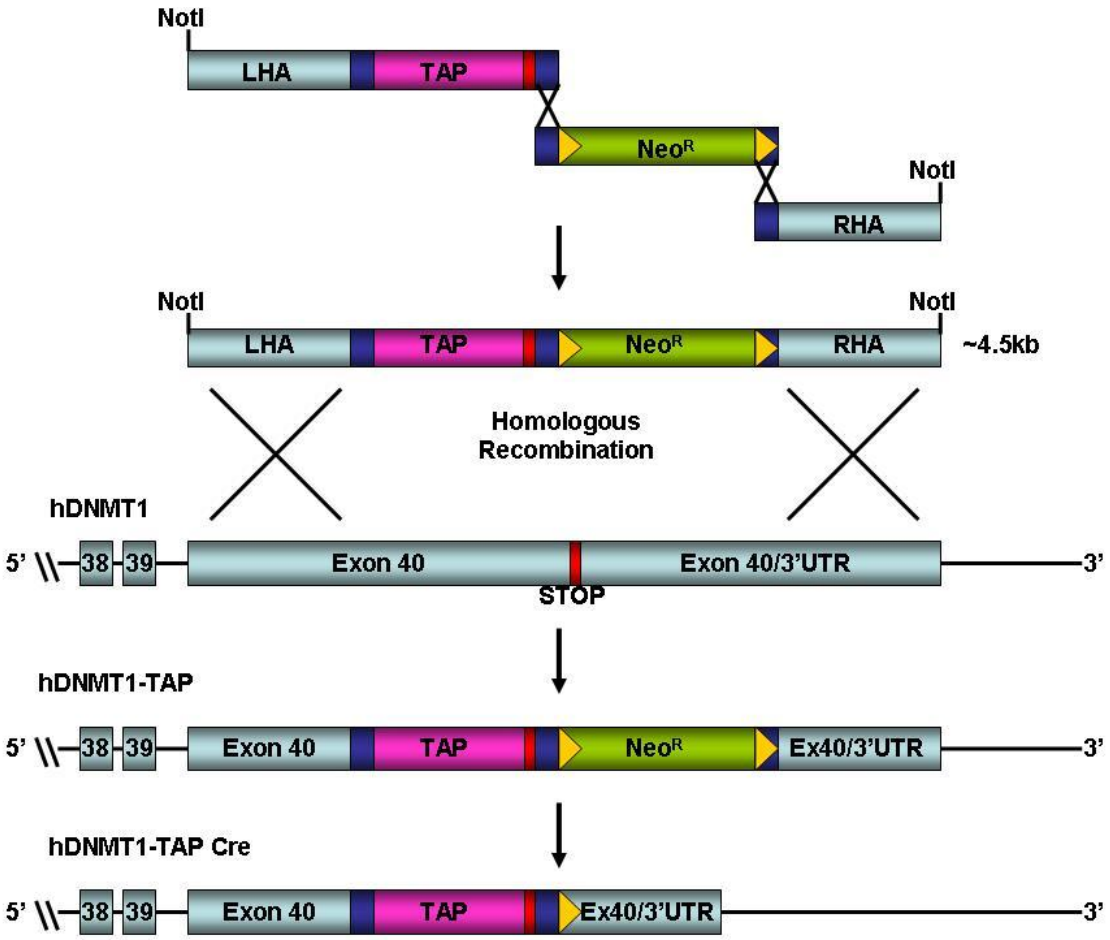


Figure 3-4: Generation of the HCT116 TAP-tagged DNMT1 cell line. The final DNMT1-TAP targeting vector was inserted into the DNMT genomic locus by homologous recombination. Neo-sensitive HCT116 parental cells were used for targeting. The original hDNMT1-TAP cell lines were selected by Neo-resistance and screened by PCR to verify proper integration of the TAP-tag. Positive recombinants were transfected with Cre-recombinase to excise the the *loxP*-floxed Neo^R cassette, resulting in the final HCT116-TAP Cre cell lines.

added to a T-25 flask, to be maintained as a back-up culture. The 96-well plates were stacked and incubated at 37°C in 5% CO_2 for approximately two weeks.

After two weeks of incubation time, the targeted HCT116-TAP cells were screened for growth of positively-recombined colonies. The stacks of 96-well plates were removed from the incubator and left to sit at room temperature for approximately 30min to allow the media to change color. Yellowing of the growth medium in wells containing growing colonies was an easy way to identify which wells were to be marked for further analysis. After approximately 30min, the plates were held up to the light and all wells with an obvious color change were marked with a "/" to indicate apparent growth in that well. Old growth medium was aspirated off from all wells, and 100µL of 1x Trypsin-EDTA was added to each marked well only. Once trypsin had been added to all appropriate wells, the lid of the plate was replaced, and the plate was shaken by hand gently to aid detachment of the cells from the plate. Using a multi-channel pipettor, 150µL of fresh RPMI 1640 growth medium, supplemented with 10% FBS, 100U/mL Penicillin, 100mg/mL Streptomycin and 400µg/mL G418 was added to each well of two new 96-well plates. These plates were labeled and set-up identically, such that each well on one plate corresponded to a well in the same position on the other plate. An aliquot (50µL) of each trypsinized clone was transferred into the same corresponding wells on two identical 96-well plates. One plate was labeled to be harvested for purification of gDNA, the other plate was labeled to be frozen as a back-up.

PCR screening to verify correct integration of the TAP-tag

To one set of duplicate 96-well plates, 50µL of freezing medium (FBS and 10% dimethyl sulfoxide (DMSO)) was added to each well. Plates were sealed individually with parafilm (Thomas Scientific, cat # 7315D45) to prevent evaporation, and stored at -80°C as a back-up.

The other set of duplicate 96-well plates was emptied of all growth medium, wrapped in parafilm and frozen at -80°C for storage until DNA purification. Total genomic DNA (gDNA) was isolated from 24 of the positively-recombined, G418-resistant HCT116-TAP clones from only one plate first, using the Gentra Puregene Cell kit (Qiagen, cat # 158388). The recommended protocol for DNA purification from cultured cells was followed as described in the manufacturer's instructions, however, because the number of cells plated in the 96-well format is so low, the volumes used for each reagent were scaled-down proportionately. DNA purification was performed using the following volumes: 60µL of Cell Lysis Solution, 2µg of RNase A, 20µL of Protein Precipitation Solution, 60µL of Isopropanol, 60µL of 70% ethanol, and 10µL DNA Hydration Solution. DNA dissolved in DNA Hydration Solution was incubated at room temperature with gentle end-over-end rotation overnight. After solubilizing overnight, sample tubes were quickspun to collect all contents at the bottom of each tube, and quantitated on the NanoDrop spectrophotometer. The amount of DNA recovered was below the limit of detection for the NanoDrop spectrophotometer. Nonetheless, we performed a "practice PCR screen" using the DNA purified from four individual G418-resistant clones (A1, A2, B1, and B2) and the P1sn-P2asn primer set, which consistently gave robust PCR products under its optimal conditions. Both a no-template negative control and a positive control reaction using HCT116 gDNA were included. The amplification products from this "practice PCR" were resolved on a 1% TAE-EtBr agarose gel, and showed bands of the expected 950bp size in all sample lanes, except the negative control. Thus, although we could not get an accurate quantitation of the amount of DNA purified from the 96-well plate, the results of this practice run gave us confidence that we had enough to proceed with screening by PCR, and that the purified DNA was of sufficient quality to yield robust PCR products.

For each row of positive recombinant DNMT1-TAP clones (A1-12, and B1-12), a large PCR mastermix in a total volume of 300µL was prepared by combining 150µL 2x iProof GC-rich mastermix, 0.6µM primer P5a, 0.6µM primer TAPrev (Figure 3-5 and Table 3-1) and 2µL template DNA from each DNMT1-TAP clone. The tubes of PCR mastermix were pulse-vortexed and quickspun to ensure all contents were mixed thoroughly and collected at the bottom of the tubes. 24µL of PCR mastermix was aliquoted into each of 12 wells contained on 0.2mL clear 8-tube PCR strips. 2.0µL of DNA purified from each positive recombinant DNMT1-TAP clone was added individually to each of the 12 wells, and the complete reaction was stirred gently with the pipet tip to mix. The PCR cycling program consisted of initial denaturation at 94°C for 2min, followed by 35 cycles, each including: denaturation at 94°C for 30sec, annealing at 59°C for 30sec, extension at 72°C for 1min, followed by final extension at 72°C for 5min and a hold at 4°C. PCR products were resolved by gel electrophoresis on a 1% TAE-EtBr agarose gel, and visualized by UV transillumination (Figure 3-6A and 3-B LHA).

Cre recombinase-mediated excision of Neo^R cassette

The pMC1-Cre recombinase plasmid (obtained from Dr. Jolene Windle, VCU) was used to excise the bulky Neo^R cassette from the targeted DNMT1-TAP allele. Adapting protocols from both Cortney Heyer in Dr. Moran's laboratory, and the manufacturer's instructions for the Lipofectamine 2000 reagent, we developed a transfection protocol that worked best for our purpose. Briefly, we used a 6-well plate format, with 2mL total plating volume, and transfected 4.0µg of the pMC1-Cre plasmid using 10µL of Lipofectamine reagent.

The morning of transfection, the growth medium on cells to be transfected was changed to fresh RPMI 1640, supplemented with 10% FBS only (in the absence of PenStrep, and

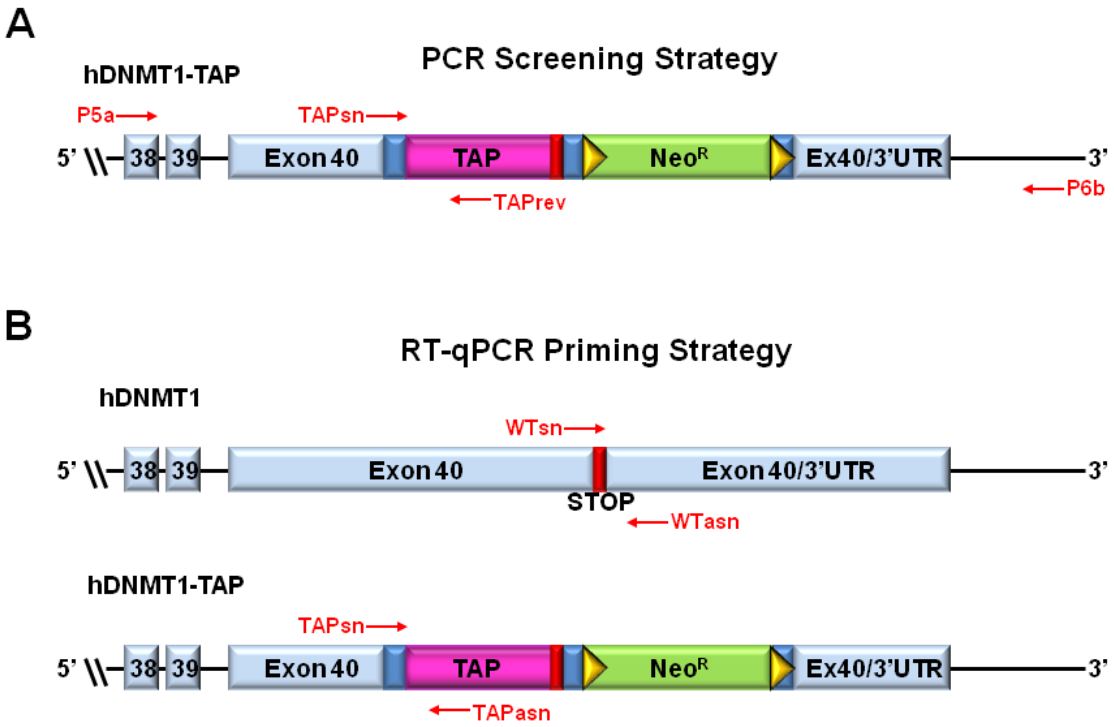


Figure 3-5: PCR screening and RT-qPCR strategies. (A) A sense primer placed across the LHA-TAP fusion junction and an antisense primer placed within the TAP-tag were used to screen G418-resistant clones for proper insertion of the TAP tag into the genomic DNMT1 locus. (B) Abundance of the WT DNMT1 allele was measured using a sense primer crossing the native termination codon for DNMT1, and an antisense primer within the 3'UTR of Exon 40. Abundance of the DNMT1-TAP allele was measured with the TAPsn primer, paired with an antisense primer placed within the TAP-tag.

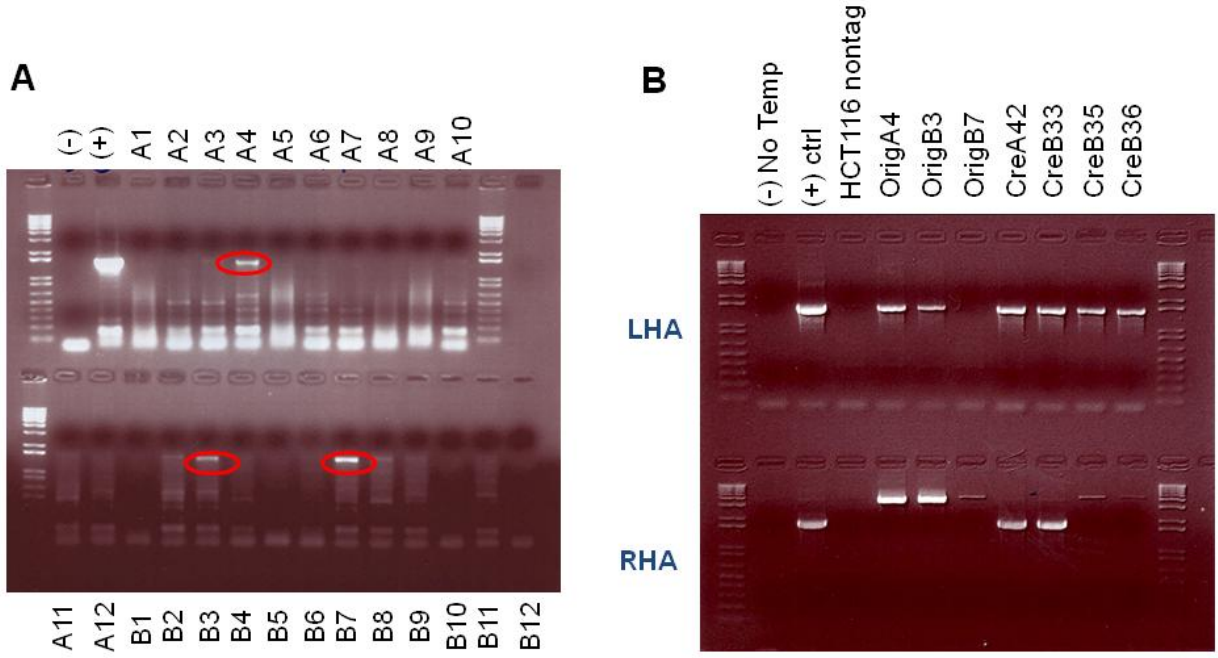


Figure 3-6: Screening of clones by PCR identifies positive recombinants. (A) Neo-resistant clones were screened first by PCR to verify proper integration of the TAP-tag. Clones A4, B3 and B7 all showed a positive result. These original clones were transfected with Cre recombinase and subcloned to ensure that all subsequent analysis was performed on homogeneous cell lines. (B) Cre Subclones and their parental lines were again screened by PCR, this time with primers directed against both the LHA and the RHA. Two TAP Cre subclones (CreA42 and CreB33) screened correctly from both sides of the insertion site. The Orig B7 clone turned out to be a mixed population of cells, and was subsequently discarded from further analysis.

400µg/mL G418). The cells were allowed to incubate in this medium for approximately 3hrs prior to transfection. Under the tissue culture hood, two sterile 1.5mL microcentrifuge tubes were labeled for each clone to be transfected. One tube was prepared with 4.0µg of the purified pMC1-Cre DNA, and brought to a total volume of 250µL with OptiMEM solution. The other tube was prepared by adding 10µL Lipofectamine and 240µL OptiMEM, for a 250µL total volume. The mixtures in each tube were allowed to incubate separately at room temperature under the hood for 5min before combining tubes, to a total of 500µL transfection reaction per clone. The DNA-Lipofectamine mixtures were incubated together for 20min at room temperature under the tissue culture hood. After 20min, the entire 500µL transfection reaction was added dropwise to each well of a 6-well plate, the lid on the plate was replaced, and the plate was swirled gently to mix. Transfected cells were incubated at 37°C overnight before subcloning.

Selection, screening and verification of final HCT116-TAP Cre cell lines

Subcloning of HCT116-TAP Cre clones

Subcloning of positively-recombined cells was performed to ensure that all final cell lines were true homogeneous clones, resulting from only a single progenitor cell, and representing only a single recombination event. Clones arising from mixed populations of cells – some correctly-targeted, some random integrants – would complicate our downstream analyses. Dr. Taylor assisted with the development of this subcloning strategy, to increase the ease of screening for single progenitor cells and to save time.

Cre-transfected HCT116-TAP (HCT116-TAPCre) cells were harvested by trypsinization, and pelleted by centrifugation. Under the tissue culture hood, a 20µL drop of fresh RPMI 1640 growth medium supplemented with 10% FBS was added to each well of two destination 24-well

tissue culture plates. An estimated cell number for the plated HCT116-TAPCre cells was used to calculate the series of dilutions necessary to achieve a seeding density of 0.25cells/well for each clone. Serial dilutions were performed in sterile 50mL or 15mL conical tubes using fresh RPMI 1640 growth medium, each time, pipetting up and down several times to achieve a single cell suspension and to mix thoroughly. The appropriate volume of cell resuspension calculated for each clone was then added to the drop of medium already in each well of the two 24-well plates. The lids were replaced on all plates, and plates were incubated at 37°C for approximately 2hrs, to allow the seeded cells to adhere to the plate. After 2hrs, the plates were examined under a light microscope for the presence of only one cell per well. Wells containing what appeared to be only one cell were marked, and an additional 0.5mL of RPMI 1640 growth medium supplemented with 10% FBS was added to the marked wells. These colonies were screened again by light microscopy after 24 and 48hrs, looking for colonies no bigger than the expected size for cells with a doubling time of approximately 24hrs. Subcloned colonies were propagated by feeding 0.5mL fresh growth medium once per week until they contained enough cells to be split between duplicate plates for selection.

Once the DNMT1-TAP Cre subclones had achieved a colony size of ~100,000 cells, or approximately 50% confluency, they were harvested by trypsinization and transferred to corresponding wells on two duplicate 24-well plates. Cells were allowed to incubate for 24hrs in normal RPMI 1640 growth medium supplemented with 10% FBS before 400µg/mL G418 selection medium was added to one of the duplicate 24-well plates, as described previously. This time, however, we were looking for cells that had become sensitive to G418 selection, as this indicates the proper excision of the Neo^R cassette. After 10 days under selection, the G418-treated plates were removed from the incubator and allowed to sit at room temperature for

approximately 30min. Media is wells containing growing cells remained yellow, while those without growth (neo^S) remained a richer red color. The sensitivity to G418 selection was confirmed for each marked well by light microscopy. Corresponding wells on the G418-negative plate were then marked, and these clones were seeded into larger T-25 flasks and propagated for further characterization.

PCR screening for correct excision of Neo^R cassette

To verify proper integration of the RHA into the DNMT1 genomic locus, and to verify excision of the Neo^R cassette after transfection with Cre recombinase, a PCR screen was again employed (Figure 3-5A). This time, a forward primer placed within the TAP-tag (TAPsn, Table 3-1) was paired with an antisense primer located ~100bp downstream of the RHA (primer P6b). Total gDNA from non-tagged HCT116neo^S cells and each TAP clone of interest (both original cell lines with the Neo^R cassette and Cre subclones without the Neo^R cassette) was isolated using the Qiagen DNeasy kit, according to manufacturer's instructions. 50ng of gDNA from each cell line was added to a 25µL PCR reaction, containing 22µL of Platinum PCR SuperMix High Fidelity (Invitrogen, cat # 12532-016), 0.6µM of primer TAPsn, and 0.6µM of primer P6b. PCR cycling conditions consisted of initial denaturation at 94°C for 2min, followed by 35 cycles, each including: denaturation at 94°C for 30sec, annealing at 60°C for 30sec, extension at 68°C for 4min, followed by final extension at 68°C for 10min and a hold at 4°C. PCR products were resolved by gel electrophoresis on a 1% TAE-EtBr agarose gel, and visualized by UV transillumination (Figure 3-6B RHA).

Southern Blot

DNA probes were designed against all four regions critical to the TAP-tagged DNMT1 cell line: the LHA, the TAP-tag, the Neo^R cassette, and the RHA. LHA, TAP and RHA probes

were amplified by PCR using primers as described in Table 3-1 (SB LHAsn2-SB LHAasn2, qPCR TAPsn- qPCR TAPasn, SB RHAsn2-SB RHAasn2). PCR products from each amplification reaction were resolved on 1% TAE-EtBr agarose gels, and visualized by UV transillumination. Bands representing PCR products of the expected size for each probe were excised with a clean razor blade and purified using the Qiagen QIAquick Gel Extraction kit, according to manufacturer's instructions. The concentrations for purified DNA probes were determined using the NanoDrop spectrophotometer. An aliquot of the Neo^R cassette, excised from the pNeDaKO-Neo vector and purified for use in fusion PCR reactions to create the DNMT1-TAP targeting vector, was used as the Neo probe.

Total gDNA was prepared from non-tagged HCT116neo^S cells, as well as each DNMT1-TAP clone of interest using the Qiagen DNEasy Blood and Tissue kit (cat # 69506) according to manufacturer's instructions. The optional RNaseA treatment was performed, as recommended for RNA-free DNA preparations. Approximately 1µg of purified gDNA was resolved on a 1% TAE-EtBr agarose gel, and visualized by UV transillumination, next to both the 1kb+ DNA ladder, and λ-*Hind*III fragments. This visual check ensured that the gDNA had no visual degradation, and was of sufficient quality to be used as input for the Southern Blot digests. Once the quality of the purified gDNA was confirmed, 25µg of gDNA from each of the 21 clones of interest was digested with 25U *Bam*HI restriction endonuclease (NEB, cat # R0136L) in 1x NEB Buffer #3 with 100µg/mL BSA. Tubes were pulse-vortexed and quickspun to ensure that all reaction contents were well-mixed and collected at the bottom of the tube. Each enzyme reaction was incubated at 37°C for approximately 1.5hrs, an additional 60U (3.0µL) of *Bam*HI enzyme was added and incubations continued at 37°C overnight, to ensure complete digestion.

Approximately 1µg (8µL) of each digest was resolved on a 0.8% TAE-EtBr agarose gel, and visualized by UV transillumination, alongside the 1kb+ DNA ladder, λ-*Hind*III fragments, and 1µg of uncut gDNA to verify digestion. Although *Bam*HI is not a frequent cutter of total gDNA, a substantial difference was observed in the smearing pattern of cut versus uncut DNA, giving us confidence that the digestion was sufficient to proceed. The remaining *Bam*HI-digested gDNA samples for each clone were extracted with Phenol/Chloroform, precipitated and DNA pellets were resuspended in 20µL of 10x NEB Buffer 4, in preparation for the second restriction digest with *Ppu*MI restriction endonuclease (NEB, cat # R0506L). The remaining 24µg of gDNA from each DNMT1-TAP Cre clone was digested with 25U (5.0µL) PpuMI enzyme in a total volume of 200µL. Reactions were incubated at 37°C for approximately 2hrs, an additional 15U (3.0µL) *Ppu*MI enzyme was added and incubation was continued at 37°C overnight, to drive the reaction to completion.

Digestion was verified by resolving 1µg (8µL) of each digest on a 0.8% TAE-EtBr agarose gel, and visualizing by UV transillumination, alongside the 1kb+ DNA ladder, λ-*Hind*III fragments, and 1µg of uncut gDNA. This digest also looks as expected, with DNA smearing evident the full length of each lane, with an equal intensity of DNA staining per lane, suggesting that the DNA has been fully digested, and that an equal amount of DNA was recovered from each sample after the first phenol/chloroform extraction. *Bam*HI/*Ppu*MI-digested gDNA samples for each clone were subjected to Phenol/Chloroform extraction and precipitation as previously described. DNA pellets were resuspended in 40µL TE buffer.

*Bam*HI/*Ppu*MI DNA fragments were resolved on a 0.8% TBE-EtBr agarose gel, interspersed with lanes loaded with the1kb+ DNA ladder. The gel was electrophoresed at 40V for 14.5hrs, and visualized by UV transillumination to ensure that sufficient resolution was

achieved. The positions of each of the 1kb+ DNA ladder fragments were marked on the gel using a 21G needle and India Ink, while holding the gel over a UV light box.

The transfer sandwich was assembled by placing the gel in the bottom of an 8x8" square glass dish, and trimming the edges to remove the meniscus. Trimming the edges ensures that full contact is achieved between the gel and the BioTrans(+) nylon transfer membrane (MP, cat # 810205), however, the wells were left intact to mark their position on the membrane after transfer. The gel was covered with approximately 100mL of 0.25N HCl, and the dish was rocked at room temperature for 20min to fully depurinate the DNA contained within the gel. The gel was rinsed once with Milli-Q distilled H2O, and incubated twice with denaturing solution (1.5M NaCl, 0.5M NaOH) and neutralization solution (1.5M NaCl, 1M Tris, pH 7.4) for 20min, rinsing the gel with Milli-Q distilled H2O in-between. The bottom of a large 13x9" rectangular glass dish was pre-wet with 10x SSC (1.5M NaCl, 150mM sodium citrate, pH 7.4), then three layers of pre-cut 3MM Chromatography paper (Whatman, cat # 3030-917) were added, one piece at a time, and bubbles were removed with a hand-held flat rolling pin. The gel was removed from Neutralization solution and flipped face-down, directly on top of the 3MM papers. Placing the gel face-down will maintain the orientation of the wells from left to right on the transfer membrane. Parafilm was used to line the edges of the gel to prevent unwanted contact between the Whatman 3MM papers around the gel, and force the transfer buffer to soak through the gel only. The nylon transfer membrane was cut to size, and pre-soaked in 10x SSC. The membrane was placed carefully on the gel, making sure to line it up correctly before laying it down. This step is critical because the position of the membrane cannot be adjusted once it has touched the gel; the transfer begins immediately. An additional three sheets of Whatman 3MM paper were cut to size and also pre-soaked in 10x SSC. These were added on top of the transfer stack one at

a time and rolled to remove any bubbles, being careful not to shift the membrane below. On top of the pre-soaked 3MM papers, a 1-inch stack of bi-fold brown paper towels was placed, and the transfer stack was topped with a slightly smaller glass dish, and two half-full reagent bottles were placed inside the top dish to weight it down. The membrane was left on the benchtop overnight to allow transfer to occur by capillary action.

After transfer, the transfer sandwich was carefully disassembled, and the flattened, dehydrated gel was placed in a glass dish with H2O and EtBr and rocked for approximately 15min to re-hydrate and stain any residual DNA within the gel. A post-transfer picture was taken of the gel using UV transillumination to visualize DNA staining. The nylon membrane was placed to dry on a piece of saran wrap, and Auto-crosslinked to fix the DNA to the membrane. The crosslinked membrane was placed in the bottom of a square plastic Tupperware container, covered with saran wrap and fitted with the container lid to prevent evaporation. 50mL of Denhardt's Complete hybridization solution (50% formamide, 5x SSC, 10% Dextran sulfate, 0.5% SDS, 1x Denhardt's solution (Kim Stratton's recipe) and 0.5mL salmon sperm DNA), pre-warmed to 42°C, was added to the Tupperware container to cover the membrane. The membrane was incubated in this "prehybridization" solution at 42°C for approximately 4hrs.

During the prehybridization period, 25ng of each of the four DNA probes (LHA, TAP, Neo, RHA) was labeled using the Random Primers DNA Labeling System (Invitrogen, cat # 18187013), according to the manufacturer's standard labeling protocol. ^{32}P-dATP and ^{32}P-dCTP deoxynucleotides were purchased from Perkin Elmer (cat # BLU512H250UC, and BLU513H250UC, respectively), and used in combination with cold dGTP and dTTP, supplied with the Random Primers kit. Labeled probes were purified over illustra ProbeQuant spin columns (GE Healthcare, cat # 28-9034-08) according to manufacturer's instructions. A 1μL

sample of each probe was counted to check the level of incorporated ^{32}P-labeled nucleotides by liquid scintillation counting. Counts of 10^8-10^9 per 1µg of DNA were desired. 10mg/mL of UltraPure salmon sperm DNA (Invitrogen, cat # 15632011) was added to each purified, labeled probe. This DNA solution was boiled for 5min, and quenched on ice, before adding to the pre-hybridization solution, and incubated at 42°C overnight.

After overnight hybridization with the ^{32}P-labeled probes, the membrane was washed as follows: two times with 5x SSC + 0.1% SDS at room temperature, once with 5x SSC + 0.1% SDS at 50°C, and twice with 2x SSC + 0.1% SDS at 50°C. The membrane was checked for wash efficiency with a Geiger counter, looking for distinguishable spots of "hot" and "cold". Washing was continued in 1-2x SSC + 0.1% SDS at 50°C until such "hot" and "cold" areas could be detected. The blot was developed by autoradiography using Genemate Blue Lite film (ISC BioExpress, cat # F-9024-8X10), an X-ray film cassette and a Konica SRX-101A developer, taking exposures of various lengths (6-72hr) until the best exposure was obtained for each probe.

Sequencing

Total gDNA from the TAP CreB33 subclone was used as a template in a series of PCR reactions to generate discrete products for sequencing. A set of three nested sense primers (LHA-TAPseq, Neo sn, P3sn) were used to sequence across the LHA-TAP fusion, the excised Neo^R cassette and the RHA, respectively. One antisense primer (P6b), placed outside the site of integration of the DNMT1-TAP targeting vector, was used to sequence across the RHA, crossing into the 3' DNMT1 genomic locus. All primer sequences are listed in Table 3-1.

DNMT1 and DNMT1-TAP Expression Analyses

Cellular lysate preparation

Tissue culture dishes containing cells plated 24hrs prior to harvest were removed from the incubator and placed on ice. Media was aspirated off, and dishes were washed twice with cold 1x PBS, pH 7.4. Cells were scraped into 2mL of 1x PBS, pH 7.4 using sterile cell scrapers. Scraped cells were collected and transferred into a pre-weighed 15mL conical tube (BD Falcon). Cells were pelleted by centrifugation at approximately 900xg (2,200rpm) for 5min at 4°C in the Beckman J6-MI centrifuge. PBS was aspirated off, and the pellet was weighed. The cell pellet was resuspended in a volume of SDS lysis buffer ten times the pellet weight. Whole cell lysates were passed through a sterile 21G needle fitted with a 1mL syringe approximately 20 times to shear the gDNA and create a homogeneous mixture. Lysates were aliquoted and stored at -80°C.

SDS-PAGE and Immunoblotting

Protein concentrations were calculated using a Bradford Protein assay, as described in Chapter 2. 40µg of total protein from each whole cell lysate was diluted in an equal volume of Laemmli sample buffer plus 5% β-mercaptoethanol, boiled for 5 minutes, and quickly spun to collect the samples. Boiled lysates were loaded onto a 4-15% gradient SDS-PAGE gel unless otherwise mentioned and resolved by electrophoresis in a Mini-PROTEAN electrophoretic box unit at 150V for approximately 1hr.

Gels were transferred onto Immobilon PVDF membranes using the wet transfer method and Transfer Buffer containing SDS. 100mL of 10x Wet transfer buffer with SDS (16.879g Tris-HCl, 17.299g Tris Base, 144.134g Glycine, 10g SDS in 1L total volume) was added to a 1L graduated cylinder, and combined with 200mL of 100% MeOH, and 700mL H2O. Transfer buffer, sponges, along with the transfer tank and assembly were pre-chilled at 4°C for

approximately 1hr prior to setting up the transfer. Six pieces of Whatman 3MM Chromatography paper and one piece of PVDF membrane were cut to the size of the gel. Once the gel had finished running, it was removed from its cassette, and rinsed in Milli-Q distilled H2O for approximately 2-3 minutes to remove residual running buffer. The pre-cut PVDF membrane was soaked in 100% MeOH for 1min to activate it, and then rinsed in Milli-Q distilled H2O for 2min. Both gel and membrane were then equilibrated in chilled 1x Wet transfer buffer with rocking at room temperature for approximately 15min. After equilibration, the transfer sandwich was assembled by placing one of the black and clear transfer plates on the benchtop such that the black plate is facing down. One sponge, soaked in 1x transfer buffer was added on top of the black plate, and rolled to remove any bubbles. Then, the first three sheets of pre-soaked Whatman 3MM paper were added on top of the wet sponge, one at a time, rolling out any bubbles after each paper is placed. The equilibrated gel was then carefully placed on top of the transfer stack, and rolled to remove bubbles. The equilibrated PVDF membrane was placed on top of the gel, ensuring that it covered the entire gel. The last three sheets of pre-soaked Whatman 3MM paper were added on top of the membrane, one at a time, being careful not to shift the membrane. The transfer stack was topped with a second sponge, also soaked in 1x transfer buffer, and the transfer plate was closed and locked shut. The transfer plate was fitted into the black and red holder, with the black side of the plate facing the black side of the holder. The holder was lowered into the transfer tank, and the tank was filled with 1x wet transfer buffer with SDS. The transfer unit was plugged into a BioRad power supply and set to run at 100V for approximately 50min.

Membranes were blocked, washed and developed as described in Chapter 2. Blots were probed with various antibodies specific to DNMT1: Abcam N-DNMT1 (cat # ab16632), Sigma

N-DNMT1 (cat # D4692), Sigma C-DNMT1 (cat # D4567) using conditions as outlined in Table 3-2. An antibody directed against the CBP domain of the TAP-tag was used to detect the DNMT1-TAP allele, referred to as simply the "TAP-tag" antibody (Open Biosystems, cat # CAB1001).

Total RNA isolation

Total RNA was isolated from non-tagged HCT116neo^{S} cells and HCT116-TAP cells using the TRIzol reagent (Invitrogen, cat # 15596-026) as recommended by the manufacturer. RNA was reverse-transcribed with the SuperScript III First-Strand Synthesis system (Invitrogen, cat # 18080051) using oligo-d(T) primers, according to manufacturer's instructions. No reverse transcriptase ("No RT") negative controls were included with each sample, and endpoint PCR was performed with β-actin primers against all cDNA template samples to validate that no PCR product could be amplified from any of the No RT negative controls before proceeding with RT-qPCR, and that a product of the appropriate size and intensity was obtained, indicating that cDNA synthesis was efficient across all samples.

RT-qPCR

RT-qPCR was performed using Quantitect SYBR Green mastermix and primers designed to distinguish the non-tagged (WT) DNMT1 allele from the TAP-tagged DNMT1 allele. WT DNMT1 primers consisted of a sense primer that crossed through the STOP codon (absent in the TAP-tagged allele), and contained 11bp of sequence specific to the WT allele, paired with an antisense primer located approximately 150bp downstream, within Exon 40 of DNMT1. TAP-tagged DNMT1 primers consisted of a sense primer that spanned the LHA-TAP fusion site, with 9bp of sequence specific to the 5' end of the TAP-tag, paired with an antisense primer located approximately 200bp downstream, within the TAP-tag coding sequence. Human β-actin was

Table 3-2: Optimal antibody conditions

Antibody	Manufacturer	Species	Blocking Buffer	Primary Dilution	Secondary Dilution	Secondary Antibody	Protein Size (kDa)
N-DNMT1	Sigma	rabbit	StartingBlock	1:1000	1:15,000	goat anti-rabbit	185
C-DNMT1	Sigma	rabbit	StartingBlock	1:1000	1:15,000	goat anti-rabbit	185
TAP	Open Biosystems	rabbit	Blotto	1:500	1:10,000	donkey anti-rabbit	varies

used as a loading and normalization control for nearly all RT-qPCR studies, and these primers amplify a product of approximately 120bp. Primers against the far C-terminal end of DNMT1 (Ex37sn and Ex38asn), common to both WT and TAP-tagged alleles, were used as an additional normalization control, to rule-out any differences in primer efficiencies between the WT and TAP primer sets. A 1:25 dilution of cDNA for each sample was read from a standard curve generated by 1:5-1:125 serial dilutions of cDNA from non-tagged HCT116 cells.

PCR reactions were run on a BioRad DNA Engine Peltier thermal cycler fitted with a Chromo4 Real-Time Fluorescence Detector attachment, and data was analyzed using OpticonMonitor3 software. The PCR cycling program consisted of initial denaturation at 95°C for 15min, followed by 40 cycles, each including: denaturation at 94°C for 30sec, annealing at the pre-determined optimal temperature for each primer set (Table 3-1) for 30sec, extension at 72°C for 30sec, followed by final extension at 72°C for 5min. A melting curve was run at the conclusion of each RT-qPCR reaction, in the range of 40-90°C, with reads every 1°C, and a hold every 1sec. Table 3-1 lists all primer sequences and optimized annealing temperatures.

Primers were used to distinguish the WT and TAP alleles (Table 3-1), and a set of internal DNMT1 primers spanning the ATG3/Ex1-2asn boundary were used as a normalization standard, to rule-out any differences in efficiencies between WT and TAP-specific primers. β-actin was used as an RNA loading control.

RNA Immunoprecipitation with α-AUF1 antibody

Immunoprecipitation and RNA isolation

In an adaptation of the protocol described by Torrisani *et al.* (92), approximately 4 x 10^6 non-tagged HCT116neo^S and HCT116-TAP cells were seeded into three 100mm tissue culture

dishes, with one dish for each of three conditions: No antibody negative control, non-specific IgG negative control, and anti-AUF1 antibody. 24hrs after plating, cells were harvested by trypsinization, and cells from all three 100mm dishes of each cell type were pooled into one 15mL conical tube. Cells were pelleted by centrifguation in the Beckman J6-MI centrifuge by spinning at 2,500rpm for 5min at 4°C. Cell pellets were resuspended in 6.2mL 1x PBS, pH 7.4, and 2mL of the cell resuspension was transferred into each of three 15mL conical tubes, one for each condition as described above. Cells were again pelleted by centrifugation, but resuspended this time in 2mL 1x PBS, pH 7.4, 2mL nuclear isolation buffer (1.28M sucrose, 40mM Tris-HCl, pH 7.5, 20mM MgCl2, 4% Triton-X) and 6mL high-performance liquid chromatography-grade (HPLC) H2O. Cell resuspensions were incubated on ice for 20min with frequent inversion to mix and thoroughly lyse cells. Nuclei were pelleted by centrifugation at 2,500xg (3,483rpm in the J6-MI centrifuge) for 15min at 4°C. Each nuclear pellet was resuspended in 1mL RIP buffer (150mM KCl, 25mM Tris, pH 7.4, 5mM EDTA, 0.5mM DTT, 0.5% NP-40, 9µg/mL Leupeptin, 9µg/mL Pepstatin, 3µg/mL Aprotinin, 1mM PMSF and 100U/mL SUPERASin RNase inhibitor (Ambion, cat # AM2694), transferred to 1.5mL eppendorf tubes, and incubated on ice for 5min. Resuspended nuclei were sonicated using the Diagenode Bioruptor water bath sonicator for 25 cycles at the highest pulse, with 5sec on, and 5sec off, to ensure complete lysis of nuclei. Nuclear lysates were clarified to remove membranes/debris by centrifugation at 16,000xg for 10min at 4°C. The nuclear supernatant was transferred to new 1.5mL eppendorf tubes for the IP.

250µL of Protein G sepharose beads (GE Healthcare, cat # 17-0618-01) were pipeted using a 2-200µL tip with the end cut off. The protein G beads were equilibrated by washing three times in 250µL RIP buffer, spinning at 1000xg for 2min at 4°C between each wash to pellet the beads. After equilibration, the beads were resuspended in 250µL RIP buffer to make a 50% bead

slurry. 40µL of 50% Protein G bead slurry, together with antibody to AUF1 (Millipore, cat # 03-111), nonspecific IgG (supplied with AUF1 antibody) or no antibody, were added to each 1.5mL tube containing clarified nuclear supernatant. The immunoprecipitation samples were incubated at 4°C in the cold room, overnight with gentle end-over-end rotation.

Fresh RIP Buffer was made without the addition of SUPERASin, to use for washing. The beads were pelleted at 2,000xg for 1min at 4°C. Supernatant/flowthrough was collected for each sample, and retained in 1.5mL tubes labeled "input". The beads were washed in 500µL RIP Buffer (without SUPERASin) a total of three times, each time pelleting the beads by centrifugation at 2,000xg for 5min at 4°C. After the last wash was removed, the beads were resuspended in 1mL of TRIzol (Invitrogen) and 2mL of TRIzol was added to input samples. Both sets of samples were stored at -80°C until RNA could be isolated.

RNA was isolated from both IP and input samples according to manufacturer's instructions included with the TRIzol reagent. All reagent volumes were scaled to the initial volume of TRIzol used to resuspend the samples. RNA from IP samples was dissolved in 18µL DEPC-treated water; RNA from input samples was dissolved in 30µL DEPC-treated water. Samples were quantitated on the NanoDrop spectrophotometer, and the integrity of 1.5µg of RNA from input samples was checked on a 1% TAE-EtBr gel.

cDNA synthesis was performed using the SuperScript III RT First-Strand Synthesis system (Invitrogen) with oligo-d(T) primers, according to manufacturer's instructions. A no reverse transcriptase negative control was included for each RNA template. An endpoint PCR was performed using WTsn/WTasn primers to validate that the resultant cDNA was of sufficient quality to amplify DNMT1 by PCR, and to ensure that the primers could not generate products from the "No RT" negative control samples.

RT-qPCR

RT-qPCR was performed as described above for DNMT1 and DNMT1-TAP expression analyses, using the same primer sets and cycling conditions. However, for this set of studies, quantitation of a 1:5 dilution of cDNA for each IP sample was determined from a standard curve generated by 1:10-1:1000 serial dilutions of cDNA from the non-tagged HCT116 no antibody input sample.

TAP-tag purification from whole HCT116-TAP cells

Cortney Heyer in Dr. Moran's laboratory adapted a TAP-tag purification protocol based on several methods described in the literature (Figure 3-2) (107, 108). Approximately $500x10^6$ HCT116-TAP CreA42 cells, plated in twenty-six 150mm dishes were harvested by trypsinization, five dishes at a time. Cells were transferred and pooled into 50mL conical tubes, and pelleted by centrifugation. Cell pellets were resuspended in 10mL of 1x PBS, pH 7.4 containing complete EDTA-free protease inhibitor cocktail (Roche) twice to wash, and each time re-pelleted by centrifugation. Pellets were lysed in 4mL of Tween 20 lysis buffer (50mM HEPES, pH 7.5, 150mM NaCl, 1mM EDTA, 2.5mM EGTA, 0.1% Tween 20, 1mM PMSF, 1mM NaF, 1mM $NaVO_4$ and complete protease inhibitor cocktail), and pooled into a 15mL conical tube (8mL total volume). The cell lysate was sonicated using a Diagenode Bioruptor water bath sonicator for 25 cycles of 1 sec on, 1 sec off at the highest pulse, and incubated on ice for 30min to ensure complete lysis. A small volume (50µL) of unclarified whole cell lysate was retained and stored at -80°C. The lysate was clarified by centrifugation at 16,100xg for 15min at 4°C and the clarified supernatant was pooled into a 15mL conical tube. A small volume of

clarified supernatant (20µL) and 1/6th of the insoluble pellet were both retained and stored at -80°C.

IgG sepharose beads (GE Healthcare, cat # 17-0969-01) were pipeted using a 2-200µL tip with the very end cut off. 100µL of IgG beads were washed with 250µL Tween 20 lysis buffer three times, spinning at 250xg for 2 minutes at 4°C between washes. Equilibrated IgG beads were added to the 15mL conical tube containing the clarified cell lysate, and incubated at 4°C overnight with end-over-end rotation.

Following overnight incubation, the IgG beads were pelleted by centrifugation at 250xg for 2min at 4°C. Approximately 1mL of the unbound flowthrough fraction was retained and stored at -80°C. The IgG beads were washed three times with 500µL of Tween 20 lysis buffer, pelleting the beads at 250xg for 2min at 4°C after each wash. The supernatant from all IgG washes was collected and pooled into a 15mL conical tube and stored at -80°C. IgG beads were washed with 1x TEV buffer (3920µL HPLC H2O, 36µL of 0.1M DTT, and 175µL of 20x TEV Buffer supplied with the AcTEV protease enzyme (Invitrogen, cat # 12575015)) three times, pelleting the beads at 250xg for 2min at 4°C after each wash. The supernatant from all TEV washes was collected and pooled into a 15mL conical tube and stored at -80°C. Equilibrated beads were resuspended in 300µL of 1x TEV buffer, and 100U (10µL) of AcTEV protease was added to the tube. TEV protease cleavage was allowed to proceed for 2hrs at 4°C with end-over-end rotation, then the reaction was spiked with 50U (5µL) additional AcTEV protease and incubated for 2hrs at room temperature (the optimal temperature for the enzyme).

Calmodulin sepharose beads (GE Healthcare, cat # 17-0529-01) were pipeted using a 2-200µL tip with the very end cut off. 100µL of Calmodulin beads were washed three times with 250µL of Calmodulin binding buffer (10mM Tris-HCl, pH 8.0, 10mM β-mercaptoethanol,

150mM NaCl, 1mM magnesium acetate, 1mM imidazole, 2mM CaCl2, 0.1% NonidetP-40 and complete protease inhibitor cocktail), pelleting the beads by centrifugation at 250xg for 2min at 4°C after each wash. IgG beads from the TEV protease cleavage reaction were pelleted by centrifugation at 250xg for 2min at 4°C. The supernatant, containing DNMT1-TAP cleaved from the IgG beads, was transferred to a fresh 1.5mL microcentrifuge tube, and labeled the "TEV Eluate". The leftover IgG beads were washed twice with 300µL of Calmodulin binding buffer, and the beads were pelleted at 250xg for 2min at 4°C after each wash. The supernatant from these washes was added to the TEV Eluate tube, and a small aliquot of the TEV Eluate was retained and stored at -80°C. 1/250th volume of 1M $CaCl_2$ (2.4µL) was added to the TEV Eluate and the tube was inverted several times to mix. The TEV Eluate sample was centrifuged at 250xg for 2min at 4°C to remove any residual traces of IgG beads, and the supernatant was transferred into the microcentrifuge tube containing the equilibrated Calmodulin beads. The TEV Eluate was incubated with Calmodulin beads at 4°C overnight with end-over-end rotation, in the presence of high $CaCl_2$ to promote the calmodulin-calmodulin binding peptide (CBP) interaction.

Following overnight incubation, the Calmodulin beads were pelleted by centrifugation at 250xg for 2min at 4°C. The unbound flowthrough was transferred to a clean 1.5mL microcentrifuge tube and stored at -80°C to determine the efficiency of capture by WB. The beads were washed three times with 500µL of Calmodulin binding buffer, pelleting the beads at 250xg for 2min at 4°C after each wash. All calmodulin washes were retained and stored at -80°C. After washing, the Calmodulin beads were resuspended in 100µL of Calmodulin elution buffer (10mM Tris-HCl, pH 8.0, 10mM β-mercaptoethanol, 150mM NaCl, 1mM magnesium acetate, 1mM imidazole, 2mM EGTA, 0.1% NP-40 and complete protease inhibitor cocktail) and inverted four times to mix. The beads were incubated in Calmodulin elution buffer for 3min on

ice, then pelleted by centrifugation at 250xg for 2min at 4°C, and the supernatant was collected in a clean 1.5mL microcentrifuge tube. The last incubation/centrifugation step was repeated seven more times, to generate a total of eight 100µL elution fractions.

The leftover IgG beads and Calmodulin beads were boiled to remove any remaining bound proteins. 100µL of Laemmli sample buffer (BioRad, cat # 161-0737) plus 5% β-mercaptoethanol was added to resuspend each bead pellet. Beads were boiled for 5min, and pelleted by centrifugation at 250xg for 2min at 4°C. The supernatants were transferred to clean 1.5mL microcentrifuge tubes and stored at -80°C.

SDS-PAGE, immunoblotting and silver staining of purified DNMT1-TAP fractions

Protein concentrations for all samples retained during the purification procedure were determined by Bradford assay against standard curves generated using each buffer (i.e., Tween 20 lysis buffer, TEV elution buffer, etc.) and known concentrations of BSA. Two 4-15% gradient gels were loaded identically with fractions of the samples retained during each step in the purification protocol. Similarly, two 4-15% gradient gels (BioRad, cat # 161-1158) were loaded identically with each of the eight purified DNMT1-TAP elution fractions. 37.5µL of each elution fraction was combined with 12.5µL of 4x Laemmli sample buffer plus 5% β-mercaptoethanol, and boiled for 5min. One gel was probed with the anti-TAP-tag antibody (Open Biosystems, cat # CAB1001) as described previously. The other identical gel was processed immediately with the Silver Stain Plus kit (BioRad, cat # 161-0449), according to manufacturer's instructions.

DNMT1 Activity assay (^{3}H incorporation)

HCT116-TAP cells were harvested as described above for the TAP-tag purification. Tween 20 lysis buffer was prepared as described and supplemented with 100µg/mL RNase A (Sigma, cat # R6513). The cell lysates were sonicated for 10 cycles, instead of 25, and quantitated by Bradford assay. Recombinant human DNMT1 (cat # M0230S) and *M.Sss*I CpG methyltransferase (cat # M0226S) were purchased from New England BioLabs, and supplied with S-adenosylmethionine, 50x BSA and 10x DNA methyltransferase reaction buffer. Lambda DNA and Poly(dI-dC) substrate DNA were both obtained from Scott Rothbart in Dr. Moran's laboratory. ^{3}H-S-adenosylmethionine was purchased from MP (cat # 24051H).

25µL positive control reactions containing 0.5U of M.SssI methylase with 1µg lambda DNA substrate, or 1.0U recombinant hDNMT1 enzyme with 15ng poly(dI-dC) DNA substrate, each with 1µCi ^{3}H-S-adenosylmethionine, 0.5µL of 50x BSA, and 2.5µL of 10x reaction buffer. Similarly, increasing amounts of purified DNMT1-TAP lysate (0.05-4µg) were added into 25µL reactions, also containing 15ng poly(dI-dC) substrate, 1µCi ^{3}H-S-adenosylmethionine, 0.5µL of 50x BSA, and 2.5µL of 10x reaction buffer. No enzyme and no substrate negative controls were also included to measure background. All reactions were mixed thoroughly in 1.5mL microcentrifuge tubes by vortexing, quickspun to collect all contents at the bottom of the tube, and incubated at 37°C for 30min. The reactions were stopped on ice. Samples were spotted onto 2.5cm Whatman 3MM filter paper circles (cat # 1030-025) to bind DNA by anion exchange. The filter paper discs were placed into the bottom of a Büchner funnel fitted with a large Erlenmeyer flask and washed twice with 10mL of 0.5M $NaPO_4$, once with 2mL 70% EtOH and once with 2mL 100% EtOH under vaccum. The filters were allowed to air dry briefly on Kimwipes placed on the benchtop. Once dry, the filter papers were added to small 5mL plastic scintillation vials

along with 5mL of Safety Solve High Flash Point liquid scintillation cocktail (Research Products International, cat # 111177), and the radioactivity was counted using a liquid scintillation counter. The amount of methyltransferase activity was represented by the degree of incorporation of ^{3}H-labeled methyl groups onto the DNA substrate, and measured as counts per million (cpm).

Subcellular fractionation

Non-tagged HCT116neo^{S} cells and HCT116-TAP cells were fractionated into whole cell and mitochondrial lysates by dounce homogenization and differential centrifugation as described in Chapter 2.

SDS-PAGE and immunoblotting of subcellular fractions

Each subcellular fraction was lysed, quantitated and its constituent proteins were resolved by SDS-PAGE as described in Chapter 2. Gels were transferred to Immobilon PVDF membranes as described above, using the wet transfer protocol. An antibody directed against the far N-terminal epitope (amino acids 1-10) of DNMT1 (Abcam, cat # ab16632) was used for subcellular localization studies because it provided the strongest signal for full-length DNMT1 in the mitochondrial fraction. Antibodies against compartment-specific markers were used to demonstrate the purity of each subcellular fraction: VDAC (Sigma) as a mitochondrial marker and H3K4me3 (Upstate Biotechnology) as a nuclear marker.

RESULTS

Creation of the HCT116 DNMT1-TAP cell line by AAV-mediated gene targeting and homologous recombination

A tandem affinity purification tag was inserted into the C-terminus of a single endogenous human DNMT1 allele by adeno-associated viral (AAV) vector-mediated homologous recombination (Figure 3-4). The targeting construct contained a Neo^R cassette for selection of positively-recombined clones. Neomycin-resistant HCT116-TAP clones were screened first by PCR to check for correct insertion of the TAP-tag into the genomic locus of DNMT1. A PCR screen of the LHA showed that a total of three clones (A4, B3 and B7) produced bright bands of the correct size (~1.7kb), indicating that they contain the DNMT1-TAP targeting construct in the proper orientation, and represent products of successful homologous recombination between the pAAV-MCS-DNMT1-TAP targeting vector and the endogenous DNMT1 locus (Figure 3-6A, and 3-6B LHA). Additionally, a PCR screen of the RHA showed that a total of two original TAP clones (A4, and B3) and two Cre subclones (CreA42, and CreB33) produced bands of the correct size (~4.2kb with Neo^R cassette, and ~2kb without Neo^R cassette, respectively) (Figure 3-6B RHA). The original TAP B7 clone appeared to be a mixed population of cells (Figure 3-7), containing at least one colony representing a proper integration of the DNMT1-TAP targeting construct (hence the initial positive result by PCR screen of the LHA in Figure 3-6), and at least one colony representing perhaps a random integration of the DNMT1-TAP targeting construct (hence the failure to maintain expression of the TAP-tagged DNMT1 allele in Figure 3-7).

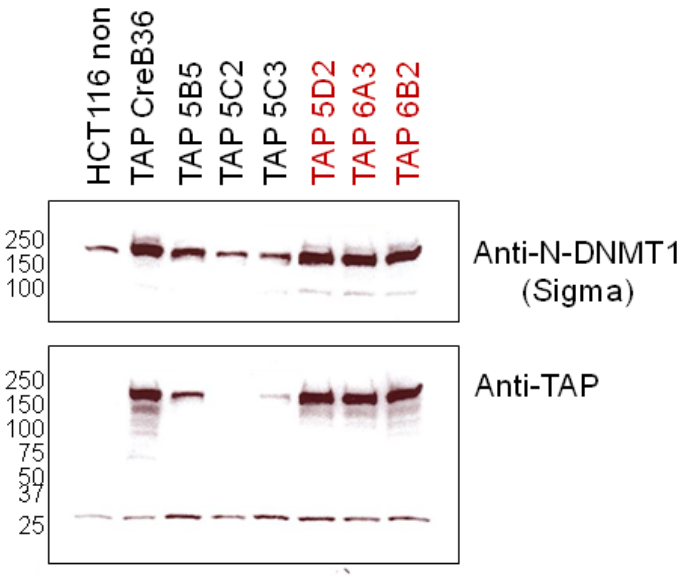


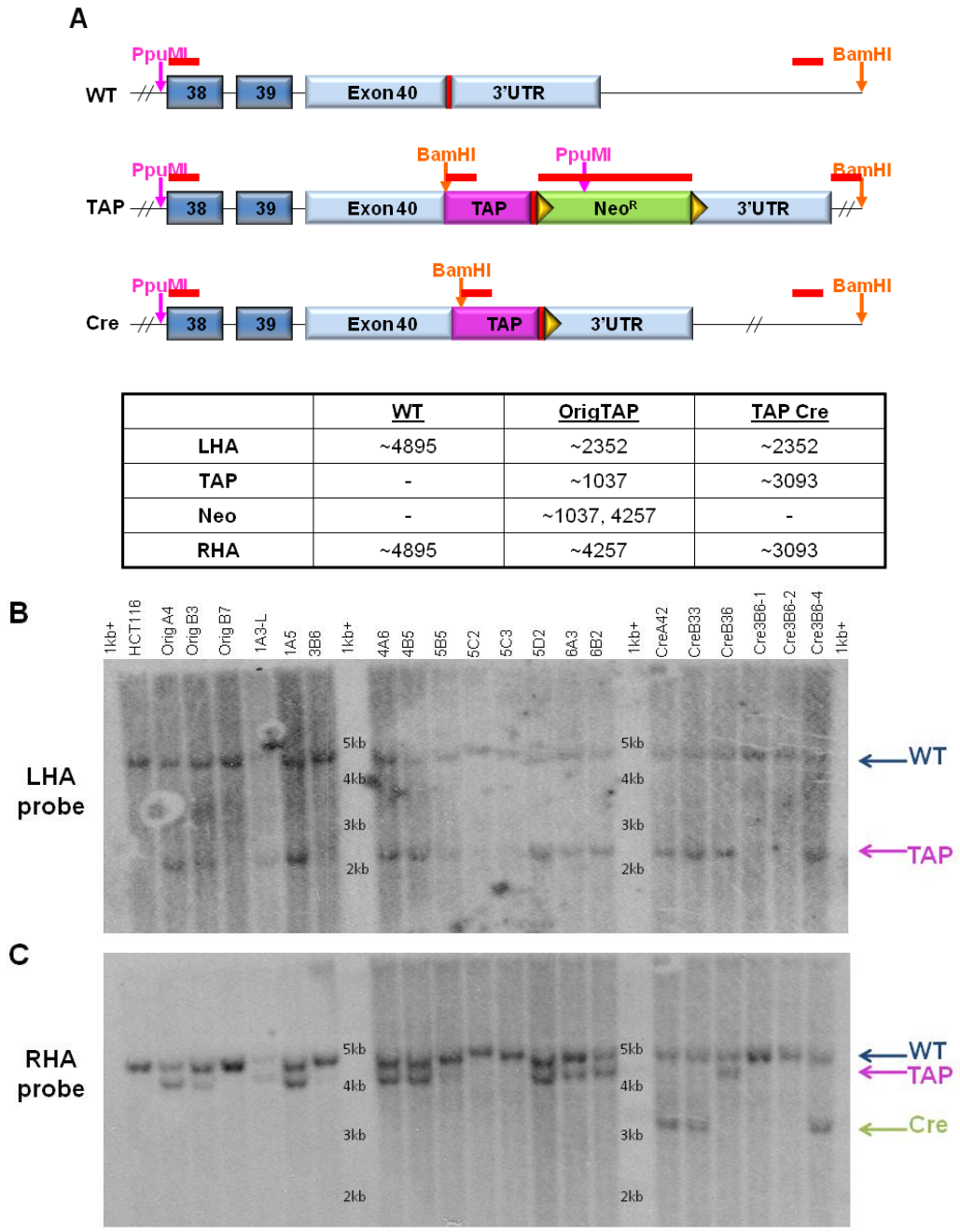
Figure 3-7: Mixed populations of Neo^{R} clones are detected by immunoblot. This is a representative immunoblot from numerous protein expression analyses performed on the HCT116-TAP clones. Whole cell lysates from each of these eight cell lines was separated by SDS-PAGE and probed with antibodies to DNMT1 and the TAP-tag. The TAP 5B5, 5C2 and 5C2 clones also represent mixed populations of cells, as it shows no expression of the TAP-tag (lower panel). The three clones labeled in red font represent true positive recombinant clones, containing proper integration of the TAP-tag in its correct orientation, and exhibiting a strong signal for TAP-tag expression (lower panel).

Subcloning of positive recombinants was required to obtain a homogeneous cell population

Because the highly efficient rAAV method for delivery of our DNMT1-TAP targeting construct produced a higher recombination frequency than expected (5.2%, see below), it became apparent that simply plating the Neo^R clones by limiting dilution was not sufficient to select for colonies arising from a single progenitor cell. Instead, many of the initial "positive" clones turned out to be mixed populations, as described above for the original TAP B7 line (Figure 3-7). An additional subcloning step was performed to ensure that any further analysis was limited to homogeneous Neo^R clones that represent a single recombination event.

Southern Blot confirms correct genotype of three final DNMT1-TAP Cre cell lines

A Southern blot with LHA and RHA probes verified the genotype of the two original TAP clones (A4 and B3), and the two TAP Cre subclones (CreA42 and CreB33), among others (Figure 3-8). A double digest with restriction endonucleases *Bam*HI and *Ppu*MI, which cut the DNMT1 genomic locus outside of the insertion sites of the DNMT1-TAP targeting vector, produced a 4.9kb fragment from the WT DNMT1 allele. *Bam*HI also cuts once within the TAP-tag, and *Ppu*MI cuts once within the Neo^R marker, so a double digest of the original DNMT1-TAP allele produced three fragments: ~1kb, ~2.4kb and ~4.2kb in size (Figure 3-8A). Because the TAP Cre clones should lack the Neo^R cassette, a double digest of the DNMT1-TAP Cre allele was expected to produce two fragments: ~2.4kb and ~3.1kb in size. The blot probed with a LHA probe (Figure 3-8B) shows a total of 13 HCT116-TAP clones with the proper fragment sizes: one band at ~2.4kb representing the DNMT1-TAP allele, and one band at ~4.9kb, representing the WT DNMT1 allele. The blot probed with a RHA probe (Figure 3-8C) shows a total of 12 HCT116-TAP clones with the proper fragment sizes: one band at ~4.9kb representing



	WT	OrigTAP	TAP Cre
LHA	~4895	~2352	~2352
TAP	-	~1037	~3093
Neo	-	~1037, 4257	-
RHA	~4895	~4257	~3093

Figure 3-8: Southern blots confirm genotype and proper integration of TAP-tag targeting construct. Total gDNA was purified from all HCT116-TAP clones of interest; both original TAP lines and cell lines that had been transfected with Cre recombinase. The DNA was digested with BamHI and PpuMI restriction endonucleases, cutting at sites marked by red arrows. Probes directed against all four components of the TAP-tag (red bars) were labeled with ^{32}P. Expected sizes of bands detected with each probe, for each type of DNMT1 allele are listed in the table. (B) A total of 13 clones produced bands of the correct sizes when visualized with a probe against the LHA. (C) Only 12 clones produced the correct fragments when probed with the RHA probe.

the WT DNMT1 allele, and either one band at ~4.2kb representing the original DNMT1-TAP allele with the Neo^R marker, or one band at ~3.1kb representing the DNMT1-TAP Cre allele without the Neo^R cassette. Technical issues, including bubbles in the transfer occluding the visualization of key bands, and non-specific binding of the Neo probe to plasmid sequences used to generate the 1kb+ DNA ladder, prevented the blots probed with TAP-tag and Neo probes from being interpreted. Overall, out of the 232 Neo-resistant subclones examined, a total of 12 HCT116-TAP clones were verified to contain the correct genotype, resulting in a recombination frequency of 5.2%.

Expression of the DNMT1-TAP protein

Immunoblots probed with two of the commercially-available DNMT1 antibodies (Sigma N-terminal, and Sigma C-terminal, see Figure 3-9A) showed a substantial difference in DNMT1 expression between the non-tagged HCT116 cells and HCT116-TAP cells (Figure 3-9B). The Sigma N-terminal antibody is epitope-specific, directed against amino acids 72-78 of the nuclear DNMT1 protein; the Sigma C-terminal antibody is also an epitope-specific antibody, directed against amino acids 1581-1595. The difference in DNMT1 protein between non-tagged HCT116 cells and HCT116-TAP cells was striking. On immunoblots probed with either of these two antibodies, DNMT1 appeared to be at least 10-20-fold more abundant in the TAP-tagged cells (Figure 3-B).

In contrast to these results, an antibody obtained from Abcam (cat # ab16632) and directed against a far N-terminal epitope of DNMT1 (amino acids 1-10), showed an equivalent level of DNMT1 expression across non-tagged and TAP-tagged cells (Figure 3-9B). Furthermore, when the WT and DNMT1-TAP alleles were resolved on a 5% SDS-PAGE gel, the

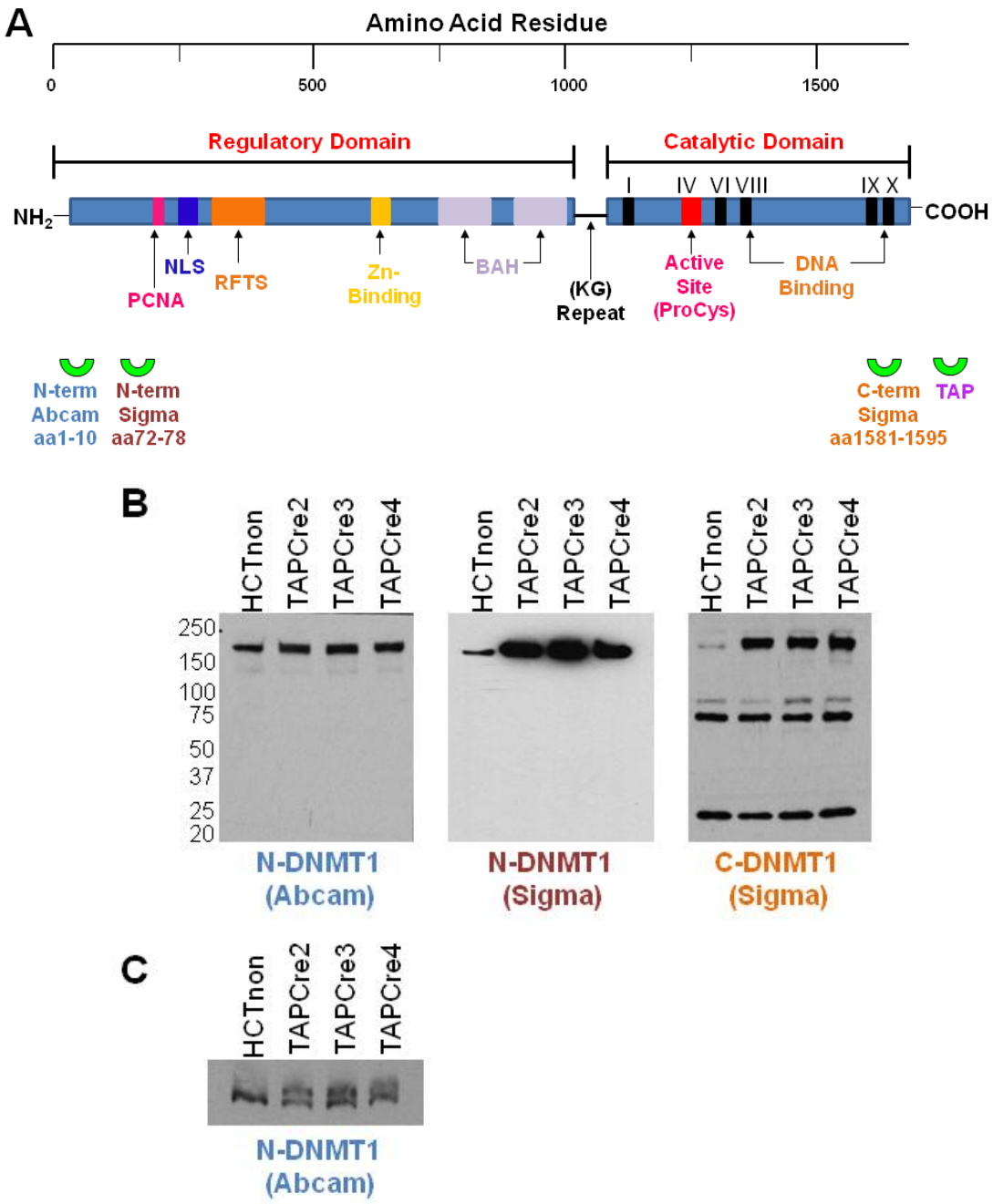


Figure 3-9: Immunoblot analyses of WT and TAP-tagged DNMT1 expression. (A) Molecular structure of DNMT1, showing the locations of conserved domains within the protein. The green half-circles represent specific epitopes detected by several antibodies against DNMT1. (B) Immunoblots of whole cell lysates from non-tagged HCT116s and three of the validated TAP Cre subclones separated on 4-15% gradient SDS-PAGE gels were probed with antibodies directed against different epitopes of the DNMT1 protein. (C) Whole cell lysates from each cell line of interest were separated on a 5% SDS-PAGE gel to resolve the 20kDa in size between the WT DNMT1 (lower band) and the DNMT1-TAP (upper band) alleles.

relative levels of DNMT1 expression between the two alleles was equal (Figure 3-9C). It is still unclear why we consistently observe two different pictures using antibodies that detect the same protein, but further studies are necessary to resolve this discrepancy, including looking at the relative level of WT and DNMT1 expression with antibodies directed against other epitopes of the protein.

Expression of the *DNMT1-TAP* allele is 6- to 8-fold higher than the endogenous WT *DNMT1* allele

RT-qPCR studies using primers specific for either the WT allele or the DNMT1-TAP allele (Figure 3-5B) and random hexamer-primed cDNA examined the relative levels of *DNMT1* mRNA expression across non-tagged HCT116 cells and three of the TAP Cre subcloned lines (Figure 3-10). An internal set of primers against DNMT1 was used as a normalization control, to account for any differences in primer efficiencies. β-actin was used as a second normalization control, to account for any variability in the amount of mRNA put into cDNA synthesis. The reported dilution values for either the WT- or TAP-specific primers were normalized first to the internal DNMT1 primer set, and then those ratios were normalized to β-actin. The average for WT DNMT1 expression in the non-tagged HCT116 cells was set to one, and all other values were adjusted accordingly. We observed no non-specific amplification with DNMT1-TAP-specific primers in the non-tagged HCT116 cells. Expression of the WT DNMT1 allele in all three TAP Cre subclones was approximately 0.5ng, which would be expected for cells containing only one WT allele. However, expression of the DNMT1-TAP allele was approximately 6-8-fold higher (3-4ng) than expression of the WT DNMT1 allele in all three TAP Cre subclones. Because only the TAP-tagged allele was affected, this result suggested that C-

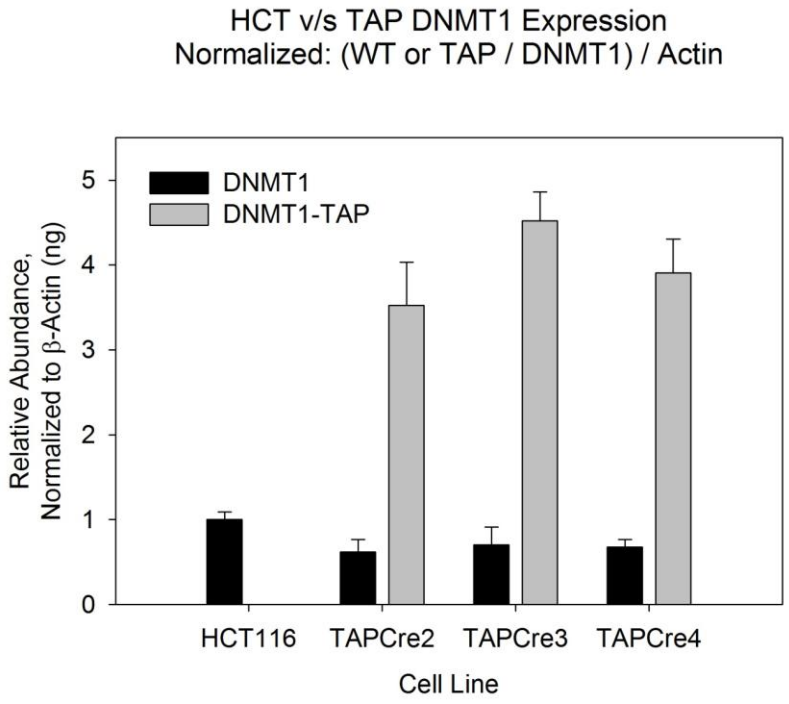


Figure 3-10: Quantitation of *DNMT1-TAP* overexpression by RT-qPCR. Total RNA was isolated from non-tagged and TAP-tagged HCT116 cells and used to synthesize cDNA with random hexamers. RT-qPCR using primers that distinquish the WT and TAP-tagged DNMT1 alleles were used to quantitate differences in expression of these two loci. Dilution values for each primer set were first normalized to internal DNMT1 primers to account for any differences in primer efficiencies. These values were then normalized to β-actin as an RNA input control. Graphed values represent the average of at least three technical repeats per sample from a representative experiment, with error bars signifying the standard deviation about the mean, calculated using the formula described in Chapter 2.

terminal insertion of the TAP-tag into the DNMT1 genomic locus may have disturbed a normal regulatory mechanism, affecting mRNA stability.

AUF1 binding and destabilization of *DNMT1* is reduced in the TAP-tagged cells, contributing to increased stability of the *DNMT1-TAP* allele

Precedent literature identified a conserved AU-rich element in the 3'UTR of *DNMT1* that is bound by AU-rich element/poly(U)-binding/degradation factor 1 (AUF1), subsequently targeting the transcript for destabilization and degradation by the exosome (92). This 54 nucleotide AU-rich element is located just 200bp downstream of the insertion site for the DNMT1-TAP targeting construct (Figure 3-11), so it stood to reason that insertion of the TAP-tag at the C-terminus of DNMT1 may have disrupted the AUF1 interaction, and caused an increase in stability of the *DNMT1-TAP* message. We therefore carried out RNA immunoprecipitation studies to assess the binding of AUF1 to the 3'UTR of the two *DNMT1* transcripts. RT-qPCR demonstrated a 50% reduction in the level of AUF1 bound to *DNMT1* mRNA in the HCT116-TAP cells, relative to non-tagged HCT116 cells (Figure 3-12). A *β-actin* negative control represents the level of background signal detected for these experiments. The observed decrease in AUF1 binding in the HCT116-TAP cells is consistent with a failure to destabilize the *DNMT1-TAP* mRNA, resulting in a higher steady-state level of the *DNMT1-TAP* transcript. While not quantitatively proportional with the level of *DNMT1-TAP* mRNA over-expression (we might expect a 6-8-fold reduction in AUF1 binding), this observation also likely explains the over-expression of DNMT1-TAP seen at the protein level.

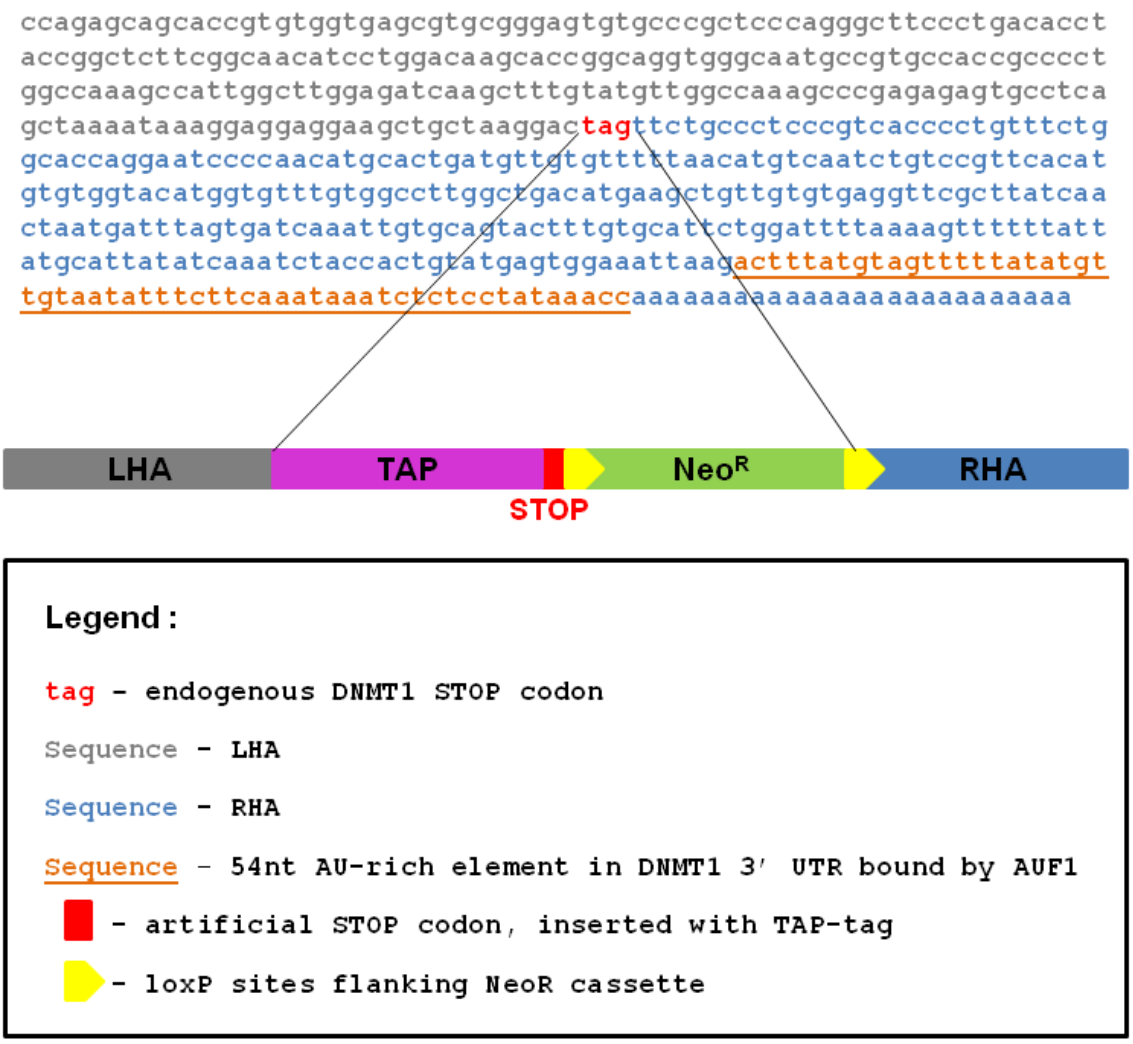


Figure 3-11: Schematic diagram of DNMT1 3'UTR, site of TAP-tag insertion, and conserved AUF1 binding site. A 54-nucleotide highly-conserved AU-rich element lies just 200bp downstream of the targeted insertion site for the DNMT1-TAP targeting vector. This AU-rich element has been shown to be bound and destabilized by AUF1.

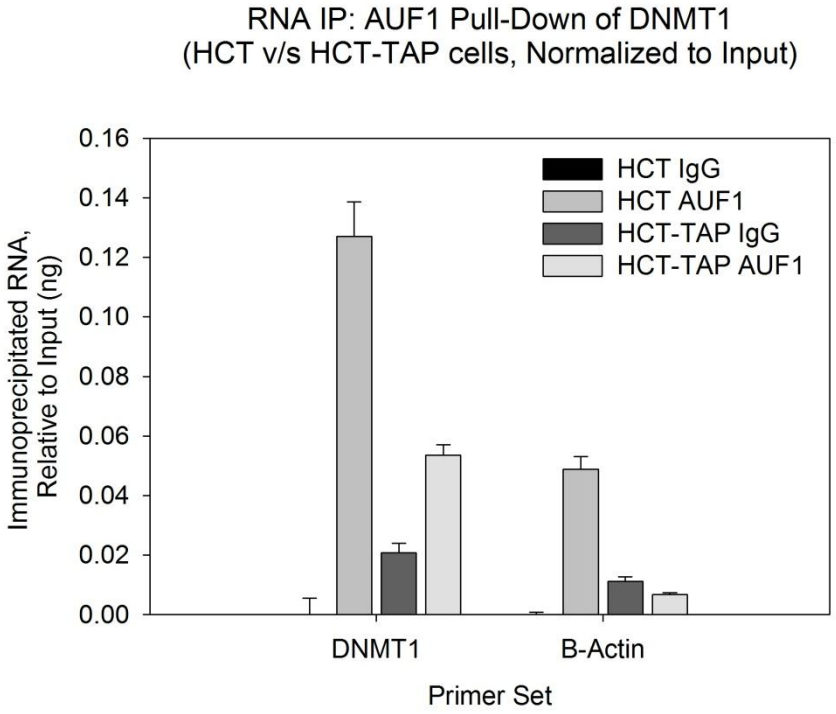


Figure 3-12: AUF1 binding to its conserved site within the DNMT1 3'UTR is disrupted in the TAP-tagged cells. RNA immunoprecipitation with an AUF1 antibody was performed in non-tagged HCT116 and HCT116-TAP cells. Nonspecific rabbit IgG was used as a negative control for the IP. Immunoprecipitated RNA was used as a template for cDNA synthesis with oligo d(T) primers. RT-qPCR using primers against DNMT1 demonstrate that the amount of *DNMT1* mRNA pulled down with AUF1 is reduced by half in the HCT116-TAP cells. Primers against β-actin were used as a negative control for RT-qPCR. Graphed values represent the average of three replicates per sample, with error bars signifying the standard deviation about the mean, calculated using the formula in Chapter 2.

Purification of DNMT1-TAP isolates multiple protein species for DNMT1

Whole cell lysates from HCT116-TAP cells were subjected to tandem affinity purification, as depicted in Figure 3-2. This procedure involved two affinity binding steps and a TEV protease cleavage step, and allowed for rapid and efficient purification of the DNMT1-TAP protein. We separated fractions of the DNMT1-TAP purification elutions on SDS-PAGE gels, and either visualized all proteins present by silver stain (Figure 3-13, top panels), or probed with specific antibodies by immunoblot (Figure 3-13, bottoms panels). As expected, the silver stained gel showed numerous bands, representing purified DNMT1-TAP, as well as any interacting protein partners that co-purified with DNMT1 (Figure 3-13B, top panel). Interestingly, when an identical gel was probed with an antibody directed against the CBP-domain of the TAP-tag (Figure 3-13B, lower panel), a similar multiple-banding pattern was observed. This suggests that not only do multiple protein species exist for DNMT1 in an endogenous context, but also that we had TAP-tagged and purified several different protein species for DNMT1.

This phenomenon was reproduced in Figure 3-14, when aliquots from several different elution fractions were pooled, separated by SDS-PAGE and probed with antibodies detecting either the extreme N- or C-terminus of the DNMT1 protein (Figure 3-9A). The TAP-tag antibody showed a familiar picture, with several protein species detected in the ~60-190kDa range. The most predominant band migrated at approximately 190kDa, representing the full-length DNMT1 (185kDa) with a fused CBP (5kDa), the part of the TAP-tag remaining after TEV protease cleavage. Likewise, an antibody against a far C-terminal epitope of DNMT1 (amino acids 1581-1595) detected multiple proteins; many, although not all, appear to correlate with bands detected with the anti-TAP-tag antibody. However, purified DNMT1-TAP probed with an antibody directed against an N-terminal epitope of DNMT1 (amino acids 72-78) only showed the

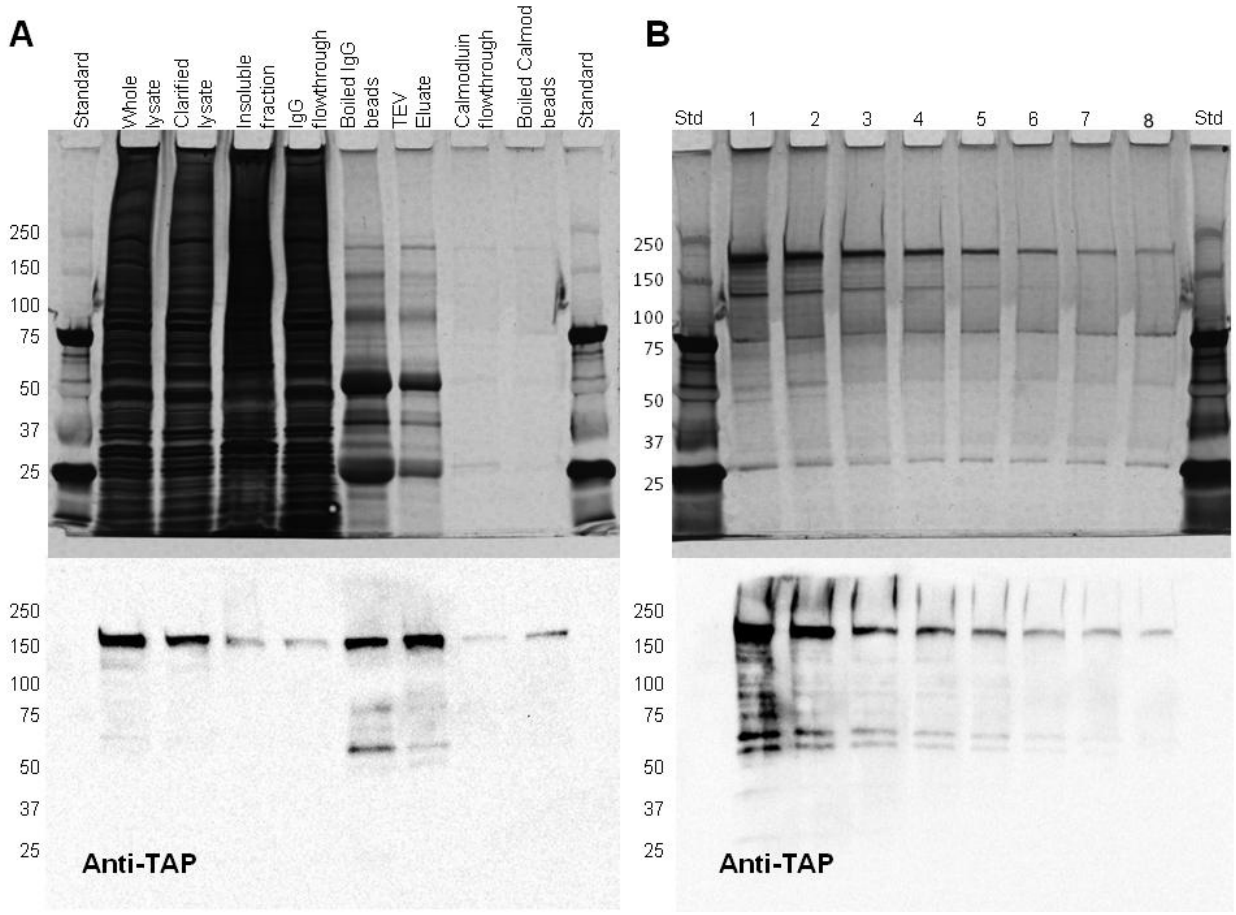


Figure 3-13: Silver stain and immunoblot analysis of the TAP-tag purified DNMT1-TAPCreA42 subclone. (A) TAP-tag purification was performed on the HCT116-TAP CreA42 cell line, samples from each purification fraction were collected and proteins were separated on two identical 4-15% gradient gels. One gel was processed with silver stain (top panel), the other was analyzed by immunoblot using an antibody against the TAP-tag. (B) Samples from each of 8 elution fractions collected from calmodulin beads were separated on two identical 4-15% gradient gels and processed as in (A).

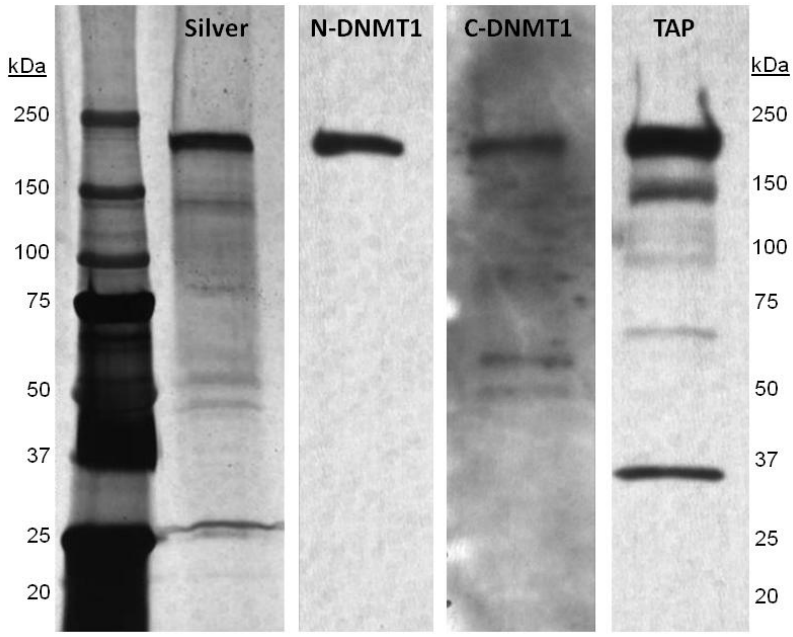


Figure 3-14: Multiple protein species are detected for purified DNMT1-TAP. Elution fractions collected from a similar TAP-tag purification as in Figure 3-12 were pooled, resolved on 4-15% SDS-PAGE gradient gels and silver stained (far lefthand panel), or probed with various antibodies against DNMT1 (three righthand panels). A similar multiple protein-banding pattern was observed in the two righthand panels, using antibodies against the C-terminus of the DNMT1-TAP allele, suggesting that multiple protein species for DNMT1 were purified by the TAP-tag purification method. When probed with an antibody against a more N-terminal epitope of DNMT1, only a single band, representing the full-length DNMT1 protein appears, suggesting that only full-length DNMT1 contains both the TAP-tag and an N-terminal epitope (aa72-78) detected by the antibody.

presence of full-length DNMT1-TAP (~190kDa). This suggests that the multiple bands detected in the two righthand panels represent species of DNMT1 lacking the N-terminus, and may arise as a result of alternative downstream start sites for DNMT1, or alternative splicing or proteolytic processing of DNMT1.

Purified DNMT1-TAP is functional as a methyltransferase

Because the entire TAP-tag purification is carried-out under mild, non-denaturing conditions, it is generally assumed that both the functionality of the target protein and its protein-protein interactions are preserved. In order to be sure that the purified DNMT1-TAP protein was functional and had retained methyltransferase activity, we performed an *in vitro* DNA methyltransferase enzyme assay, based on established methods (118, 119), and using a protocol adapted from Scott Rothbart in Dr. Moran's laboratory. The bacterial CpG methyltransferase, *M.Sss*I, in a reaction with its preferred substrate, lambda DNA, was used as a positive control. Increasing amounts of purified DNMT1-TAP protein (0.05-4µg) were added to reactions containing an unmethylated poly(dI-dC) substrate (Figure 3-15). ^{3}H-S-adenosylmethionine was used as the methyl group donor in all reactions, allowing a specific labeling of ^{3}H-methyl groups incorporated into DNA. The DNA was subsequently purified and incorporation of ^{3}H was measured by scintillation counting, as a direct measurement of DNA methyltransferase activity. Background counts were obtained from no enzyme and no template negative control reactions and an average background count was subtracted from the values reported for each enzyme. As expected, we observed a relatively high level of ^{3}H incorporation into purified lambda DNA substrate, catalyzed by *M.Sss*I. Also as expected, we observed an approximately linear increase in methyltransferase activity with increasing amounts of purified DNMT1-TAP added to the

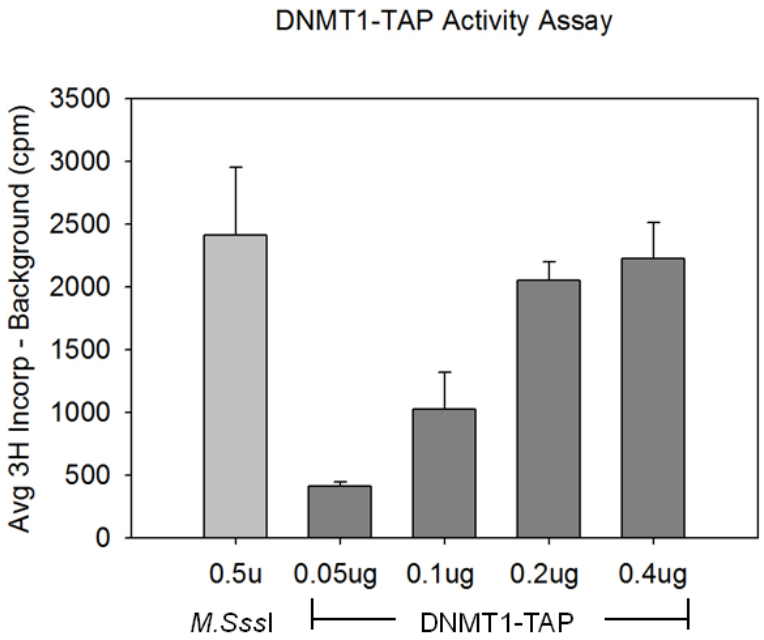


Figure 3-15: Purified DNMT1-TAP is functional as a methyltransferase. A classical *in vitro* ^{3}H-incorporation assay for DNA methyltransferase activity was performed on extracts of TAP-tag purified DNMT1. *M.Sss*I was used as a positive control. Increasing amounts of purified DNMT1-TAP produced an approximately linear increase in ^{3}H-methyl groups incorporated into a defined DNA substrate, demonstrating that the purified DNMT1-TAP allele is catalytically active. Graphed values represent the average of two replicate reactions minus background from a representative experiment, with error bars signifying the standard deviation about the mean.

reaction (Figure 3-15). Thus, the purified DNMT1-TAP enzyme is catalytically active, and capable of methylating a DNA substrate *in vitro*.

The DNMT1-TAP allele localizes properly to the mitochondria and is expressed at near-endogenous levels

To validate our use of the HCT116-TAP cell line in studies involving mitochondrial methylation, we needed to be sure that the DNMT1-TAP allele could indeed translocate to the mitochondria. Both non-tagged HCT116 and HCT116-TAP cells were fractionated into whole cell and mitochondrial extracts, and resolved by SDS-PAGE. Gels were loaded with different amounts of total protein from whole cell and mitochondrial fractions, representing equal cell equivalents. An N-terminal DNMT1 antibody (aa1-10, Figure 3-9A) shows that the level of DNMT1 in the mitochondrial fraction is approximately the same across non-tagged HCT116 cells and HCT116 cells carrying a TAP-tagged DNMT1 allele (Figure 3-16A). The appearance of a doublet in the mitochondrial fraction from the TAPCre3 cells represents the 20kDa difference in size between the WT (lower band) and TAP-tagged (upper band) DNMT1 alleles that cannot be fully resolved on a 4-15% gradient gel (Figure 3-16A). VDAC was used as both a mitochondrial marker and a loading control, and H3K4me3 was used as a nuclear marker to demonstrate the purity of these mitochondrial preparations (Figure 3-16A).

Identical immunoblots of whole cell and mitochondrial fractions from non-tagged HCT116 and HCT116-TAP cells were also probed with the Sigma N-terminal and Sigma C-terminal antibodies (Figure 3-16B). Although full-length DNMT1 could not be detected in the mitochondrial fractions of cells probed with either antibody, two smaller protein species, which we suspect are proteolytic cleavage products of the full-length enzyme, were detected. These

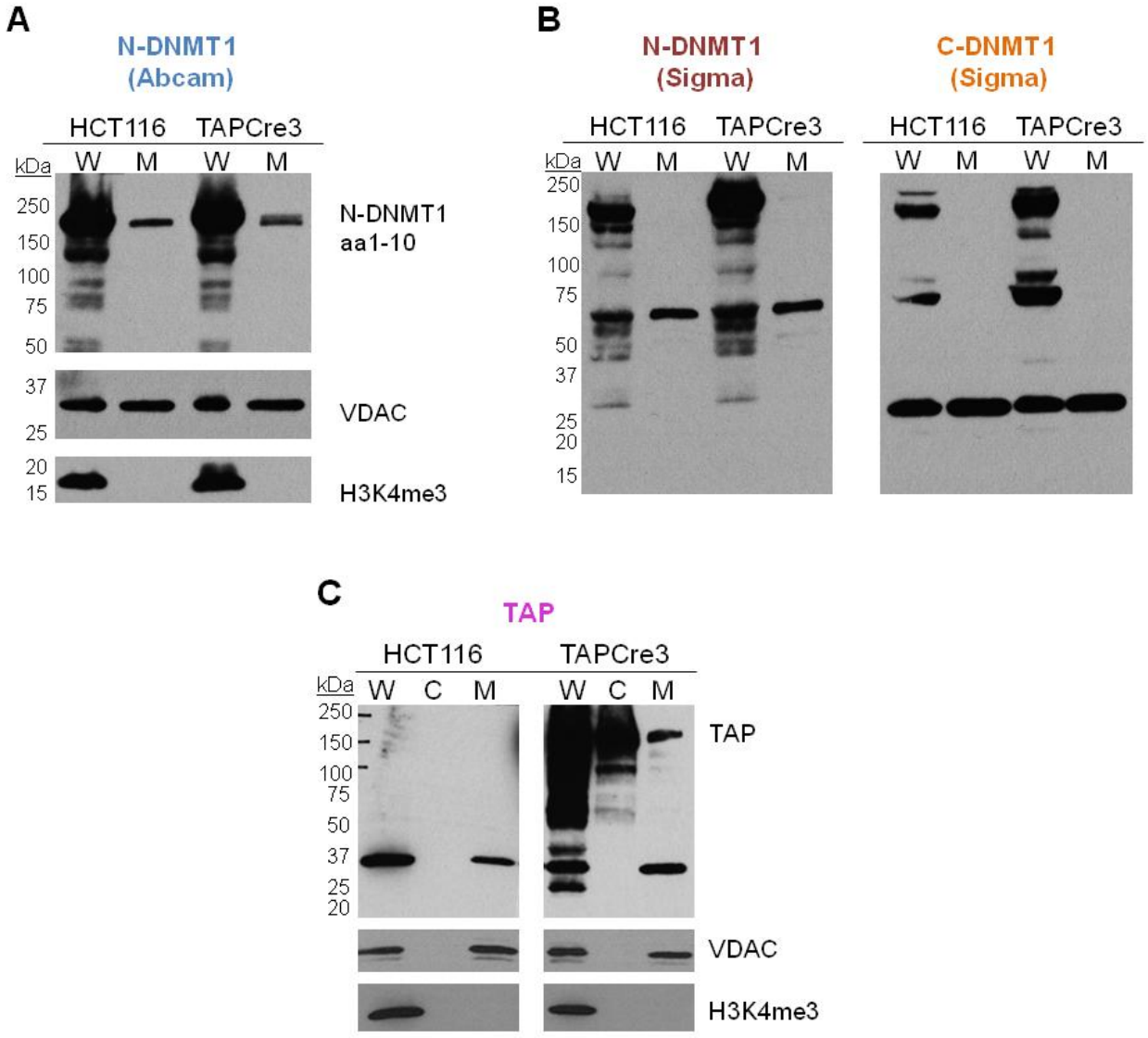


Figure 3-16: Subcellular localization of DNMT1-TAP. Non-tagged HCT116 cells and HCT116 TAPCre3 cells were fractionated into whole cell and mitochondrial lysates. These extracts were resolved on 4-15% SDS-PAGE gels, and probed with four of the DNMT1-specific antibodies (See Figure 3-9A for epitopes). (A, B) Both full-length DNMT1 and smaller immunoreactive protein species appear to be equally expressed in the mitochondrial fractions of both cell types. (C) Full-length DNMT1 along with a smaller protein species are detected in the mitochondrial fraction of HCT116-TAP cells. Subcellular fractions from non-tagged HCT116 cells probed with the TAP-tag antibody show non-specific detection of a protein of ~35kDa.

proteins (~60kDa, seen with the N-terminal antibody, and ~30kDa with the C-terminal antibody) were equally expressed across both cell lines (Figure 3-16B).

Whole, cytosolic and mitochondrial lysates from HCT116 and HCT116-TAP cells were separated on 4-15% gradient SDS-PAGE gels and probed with an antibody to the CBP-domain of the TAP-tag (Figure 3-16C). Full length DNMT1 was detected in the mitochondrial fraction of HCT116-TAP cells, validating our use of this cell line for mitochondrial studies. A smaller, ~30kDa protein band was also visible in the mitochondrial fraction, but because a band of similar size was also detected in the mitochondria of non-tagged HCT116 cells, it likely represents either a non-specific protein or an endogenous calmodulin binding protein, several of which have been detected in mitochondria purified from rat liver (120).

Consistent with the immunoblot data, RT-qPCR studies measured the abundance of mtDNMT1 using primers that distinguish the mitochondrial DNMT1 transcript from total DNMT1 (Figure 3-17). Primers against human β-actin were used as a normalization control. We observed no significant difference in the level of mtDNMT1 mRNA between non-tagged HCT116 and HCT116-TAP cells.

DISCUSSION

Strategy and rationale behind generation of the HCT116 TAP-tagged DNMT1 cell line

The TAP-tagged DNMT1 cell line was originally created as a way to track the localization and protein-protein interactions of DNMT1 within its respective subcellular compartments. None of the commercially-available antibodies against DNMT1 gave a clear picture as to where this protein partitioned within the cell, so having a way to tag and isolate

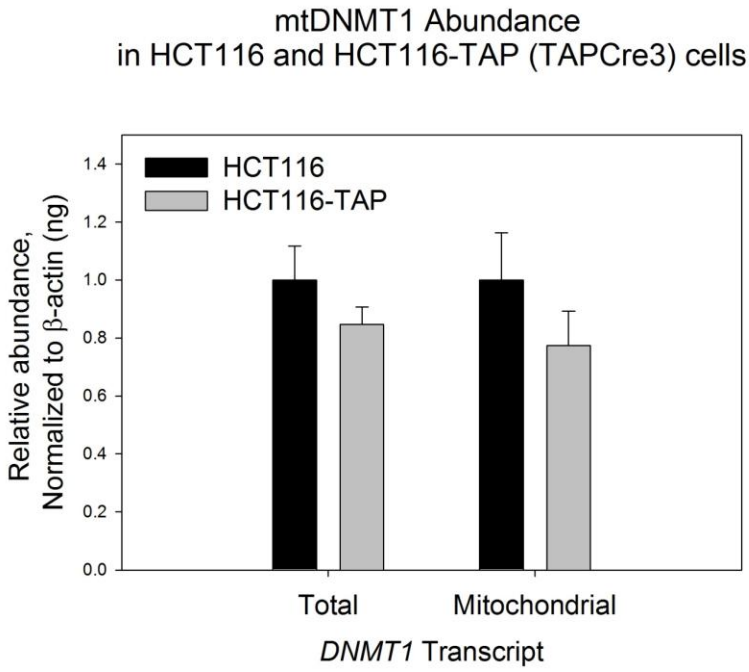


Figure 3-17: Quantitation of mtDNMT1-TAP abundance. Total RNA isolated from non-tagged HCT116 and HCT116-TAPCre3 cells was used to synthesize cDNA using random hexamers. RT-qPCR was performed using primers that distinguish the mitochondrial DNMT1 transcript from total DNMT1, as described in Chapter 2. Values were normalized to β-actin, and represent the average of triplicate samples from a representative experiment, with error bars signifying the standard deviation about the mean.

DNMT1 under native conditions became critical for our studies. Additionally, because the discovery of the mitochondrial isoform of DNMT1 lacked precedent, we needed a way to isolate mtDNMT1 interacting proteins without any prior knowledge of their identity. On the cutting edge of research at the time was the use of adeno-associated viral vector technology to generate human somatic cell knock-outs (115). Leading scientists in the field were reporting higher targeting frequencies by homologous recombination than ever achieved before. Rather than knock-out DNMT1, we decided to knock-in a clever configuration of a fusion tag that would allow us to isolate DNMT1 from is subcellular compartments, in complex with its native binding partners.

The C-terminus of DNMT1 was chosen for targeting because the precise details of the transcriptional start site for the mitochondrial targeting sequence upstream of the coding region for DNMT1 had not been mapped. Furthermore, a C-terminal tag would allow expression of the DNMT1-TAP fusion protein from its endogenous promoter, and presumably, maintain natural levels of expression of the tagged allele. As an extra precaution, the targeting construct was designed to contain a *lox*P-floxed Neo^R selection cassette, so that if it became desirable to remove the Neo^R marker, it could be easily excised by transfection with a plasmid expressing Cre recombinase.

Unexpected disruption of AUF1 binding and destabilization of the TAP-tagged *DNMT1* mRNA

Studies aimed at understanding the level of expression of DNMT1-TAP relative to WT DNMT1 generated some surprising results. Both RT-qPCR analyses of *DNMT1* expression at the mRNA level (Figure 3-10) and immunoblot analyses of DNMT1 protein levels with two

different antibodies (Sigma N-terminal DNMT1, Sigma C-terminal DNMT1, Figure 3-9) seemed to agree that the TAP-tagged DNMT1 allele was highly over-expressed relative to the endogenous DNMT1 allele. This over-expression was observed despite the fact that the DNMT1-TAP targeting construct was inserted into the genomic DNMT1 locus so that expression would be driven by the native DNMT1 promoter and the targeted allele would be subject to regulation by the same mechanisms as the WT allele. The WT allele in the HCT116-TAP cells was expressed at approximately half the level of DNMT1 in the non-tagged HCT116 cells (Figure 3-10), as expected for only one copy of the WT locus (one copy WT DNMT1, one copy DNMT1-TAP). The data suggested that manipulation of the targeted allele was the cause of this over-expression. Several reports in the literature described a conserved AU-rich element located within the 3'UTR of the *DNMT1* mRNA that could be bound by a highly-characterized destablilizing protein, AUF1 (92, 116, 117). When bound by AUF1, the DNMT1 transcript was targeted for destabilization by the exosome, and turned over rapidly (92). We hypothesized that, by insertion of the TAP-tag just 200bp upstream of the AUF1 binding site (Figure 3-11), we had introduced sufficient sequence or secondary structure to disrupt the AUF1-*DNMT1* interaction. Our RNA IP data shows substantial reduction in the association of AUF1 with *DNMT1* mRNA in HCT116-TAP cells (Figure 3-12), suggesting that, by preventing its destabilization, we induced an accumulation of the *DNMT1-TAP* transcript, consistent with our RT-qPCR results, and this increased stabilization resulted in over-expression of the DNMT1-TAP protein, as evidenced by our immunoblots.

Snapshots of purified DNMT1-TAP and speculations about other DNMT1 isoforms

In our hands, immunoblots probed with any of the commercially-available DNMT1 antibodies reproducibly demonstrated multiple protein species for DNMT1. Erica Peterson and Dr. Dolores Arjona, former members of our laboratory, performed numerous experiments aimed at revealing the true identity of these proteins, and to understand how they changed under various experimental conditions. We had hoped that, upon generation of a cell line carrying an endogenously-tagged DNMT1 allele, some of those lingering questions could finally be answered. Unfortunately, the data we have obtained using the HCT116-TAP cell line has been complicated to interpret, and in many cases, it has generated more questions than answers. What does appear to be true is that multiple proteins containing the C-terminal domain of DNMT1 are immunoreactive with several different DNMT1 antibodies; these proteins are expressed endogenously and can be purified with the TAP-tag (Figure 3-13, lower panels). Immunoblots of the purified DNMT1-TAP allele suggest that the only protein species containing both the TAP-tag and the N-terminal epitope between amino acids 72-78 is the full-length DNMT1-TAP protein (Figure 3-14). Taken together, our results are suggestive that these other, smaller proteins are N-terminally truncated versions of DNMT1, perhaps generated using downstream alternative translation start sites. We know from the data presented in Chapter 2 that an upstream in-frame start site is used to generate the mitochondrial isoform of DNMT1 (Figure 2-6). Several other ATGs in-frame with the published translational start site within the DNMT1 genomic sequence exist downstream, including ATGs at Exons 21, 25 and 34. An in-frame ATG within Exon 4 of the DNMT1 coding sequence was thought to be the primary translational start site for many years (121) until the current site was identified by 5’ RACE (93). Thus, it is not unreasonable to

hypothesize that other isoforms (122, 123) of DNMT1 may still remain to be characterized, and their discovery could provide explanation for some of our most intriguing data.

Protein binding partners predicted to have co-purified with DNMT1-TAP

Figure 3-13B (top panel) shows that numerous protein bands were detected in the purified DNMT1-TAP elution fractions on a silver stained SDS-PAGE gel. These proteins were predicted to represent multiple species/isoforms of DNMT1, as well as proteins interacting with these species of DNMT1. Sizes for the most predominant bands could be estimated using the Alpha Imager molecular weight prediction tool, and speculations were made as to the identity of these proteins, based on known interactions for DNMT1 in the nucleus (Figure 3-18). A band migrating at ~95kDa is predicted to be ubiquitin-like with PHD and ring finger domains 1 (UHRF1), a protein described to bind and tether DNMT1 to actively-replicating chromatin, assisting its faithful maintenance of the DNA methylation profile (79, 124). A protein detected at ~56kDa could be one of four known interacting partners of DNMT1 with that approximate molecular weight: histone deacetylase 1 (HDAC1), histone deacetylase 2 (HDAC2), DNA methyltransferase 1 associated protein 1 (DMAP1) or E2F transcription factor 1 (E2F1). Each of these interactions is thought to serve as a mechanism for recruitment of DNMT1 to replication foci within dividing cells, where it participates in maintenance methylation of hemi-methylated DNA (125, 126). A 27kDa protein band is predicted to represent proliferating cell nuclear antigen (PCNA), which has been shown to function as a clamp that helps to unwind DNA, and facilitate the methyltransferase activity of DNMT1 on hemi-methylated DNA (127, 128). Although supported by previous literature, these interactions are only putative; the true identities of these co-purified proteins remain to be revealed.

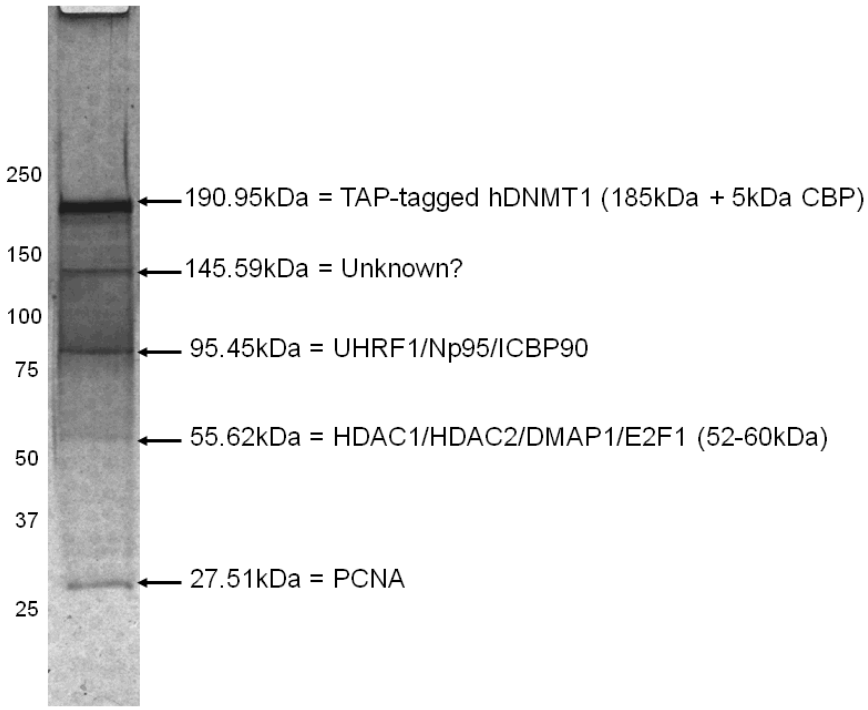


Figure 3-18: Proteins predicted to have co-purified with DNMT1-TAP. A sample of the TAP-tag purified elution fractions processed by silver stain was used to estimate the molecular weight of proteins that co-purified in complex with DNMT1-TAP. Predictions were made as to the identity of these proteins, based on what is known about protein-protein interactions for DNMT1 in the nucleus. The true identity of these protein partners remains to be tested by immunoblot.

Subcellular localization of DNMT1-TAP

Immunoblots of subcellular fractions from non-tagged HCT116 and HCT116-TAP cells demonstrate that TAP-tagged DNMT1 localizes to the mitochondria. Furthermore, expression of mtDNMT1-TAP approximates the level of endogenous mtDNMT1 expression, regardless of the antibody used for detection. We have hypothesized that the smaller protein species detected in the mitochondrial fractions using either of the Sigma antibodies are products of proteolytic cleavage of the full-length enzyme, processed upon entry into mitochondria (Chapter 2, Figure 2-10). It is not uncommon for nuclear-encoded mitochondrial proteins to undergo such modification upon translocation into the mitochondrial compartment (6, 7, 85). N-terminal presequences are regularly removed by proteolysis once the imported preprotein enters the mitochondrial matrix (6). Furthermore, several imported proteins have been shown to serve as substrates for additional processing by enzymes either bound to the inner membrane, or residing in the matrix (6). Thus, although the exact composition of the mitochondrial proteins detected at 60kDa (N-terminal antibody) or 30kDa (C-terminal antibody) is not known, it is not unreasonable to postulate that they could represent smaller forms of DNMT1. Alternatively, they could represent non-specific proteins that cross-react with the DNMT1 antibodies. In either case, they are equally expressed in both HCT116-TAP and non-tagged HCT116 cells.

Advantages and disadvantages of using the HCT116 TAP-tagged DNMT1 cell line

One major disadvantage of using the TAP-tag for purification of DNMT1 is the presence of the CBP domain. Protein recovery using calmodulin-coated beads, as described in the original TAP procedure for *S. cerevisiae* (107) was only 50%, while in other mammalian cells the recovery has been noted to be significantly lower (109, 110). Factors accounting for this could

include binding of endogenous calmodulin to the CBP-tagged DNMT1, thereby blocking the CBP site, and preventing its efficient binding to calmodulin-coated beads, or binding of endogenous CBP directly to calmodulin-coated beads, preventing their interaction with the tagged DNMT1. This major drawback can potentially be prevented, however, through the addition of a calcium chelator (EGTA or EDTA) in the first round of purification. This would help to strip the IgG binding buffer of any available Ca^{2+} that would allow an unwanted CBP-calmodulin interaction, and the endogenous calmodulin could be washed away and discarded prior to incubation with calmodulin beads.

Another significant disadvantage for any tagged protein is that the tag itself may disrupt complex formation, and/or potentially generate false positive interactions. By cleaving off the larger ProtA domain with TEV protease during the second purification step, this risk is diminished somewhat, but it is still an issue for consideration. In our case, however, it may be that the TAP-tag serves as more of a hindrance, limiting import of the DNMT1-TAP protein through the translocase machinery and into the mitochondrial compartment. In immunoblots where the WT mtDNMT1 and mtDNMT1-TAP alleles could be resolved, it appears as though the level of mtDNMT1-TAP protein is slightly under-represented relative to the non-tagged allele (Figure 3-16A), suggesting that the TAP-tag may prevent import of some mtDNMT1-TAP. Complete resolvation of the WT and DNMT1-TAP alleles in mitochondrial fractions was attempted several times on 5% SDS-PAGE gels without success. However, on well-resolved 4-15% gradient gels, like Figure 3-16A, a doublet could be visualized at ~185kDa, representing what we believe to be the WT and TAP-tagged DNMT1 alleles. In such instances, signal for the upper band, signifying the mtDNMT1-TAP allele, was less intense than the lower band,

representing the WT mtDNMT1 allele. Further experimentation is needed to verify this observation.

The TAP-tag was chosen for its rapid and efficient recovery of a fusion protein, and its relative ease of use in small-format purifications. Among its other advantages are that it allows for the purification of protein complexes in their native context, and because the purification is performed under mild, non-denaturing conditions, the purified protein is expected to retain activity. Since we have TAP-tagged the endogenous DNMT1 locus, we have a unique tool for characterizing the activity and protein-protein interactions that occur at endogenous levels, and not those that are created nonspecifically or unnaturally due to overexpression of cDNA or the presence of foreign vector sequences. Additionally, because we have employed an entirely eukaryotic system by expressing the DNMT1-TAP fusion protein in mammalian cells, it is subjected to normal regulatory mechanisms and post-translational modifications, which could be functionally important to DNMT1 (76).

Our data demonstrates that purification of the DNMT1-TAP allele is possible (Figures 3-13 and 3-14), and that the isolation procedure does not disrupt enzymatic activity of DNMT1 (Figure 3-15). We have also shown that the DNMT1-TAP allele translocates to the mitochondria (Figure 3-16), and is expressed at near-endogenous levels within this cellular subcompartment (Figure 3-17). These observations not only validate our use of the HCT116-TAP cell line for studies examining mitochondrial function, but they also support our use of this unique reagent to understand what other mitochondrial accessory proteins interact with mtDNMT1 at endogenous levels.

Chapter 4: DNA Binding and Enzymatic Activity of Mitochondrial DNMT1

INTRODUCTION

I. Preferential DNA binding by DNMT1 in the nucleus

DNMT1 has been shown to localize with discrete sites of new DNA synthesis, called replication foci, in cells undergoing S-phase (72, 129, 130). This association is mediated by a replication foci targeting sequence (RFTS) located within the N-terminal regulatory domain of the protein (Figure 3-9A) (72). Its role in faithfully propagating the methylation patterns in a post-replicative fashion earned DNMT1 the title of "maintenance methyltransferase".

DNMT1 exhibits a preference for hemi-methylated DNA on the range of 3-50-fold, depending on the experimental conditions, the length of the DNA substrate, the sequence context, and the preparation of the enzyme (130-132). However, DNA methylation rates do not usually correlate with DNA binding affinity or preference of the enzyme; the kinetics of the methylation reaction catalyzed by DNMT1 is highly dependent on the sequence context and structure of the substrate. For instance, a target CpG site embedded in a CG-rich context is methylated more slowly than a target CpG site within an AT-rich sequence (133). Similarly, single-stranded DNA and double-stranded DNA have remarkable effects on DNMT1 activity. A single-stranded sequence with one 5mCpG site within a CG-rich sequence can serve as a potent allosteric inhibitor of DNMT1 activity, whereas the same un-methylated single-stranded CG-rich sequence is a substrate for DNMT1, but binding affinity and methylation rates are slow (133). Double-stranded DNAs, however, can support methylation regardless of whether they are un-

methylated or hemi-methylated, but the binding preference and catalytic activity of DNMT1 is much higher on double-stranded hemi-methylated DNA substrates (74, 130-133).

At least two specific sites within DNMT1 have DNA-binding abilities: the active site, within the catalytic domain of the enzyme, and the allosteric (Zn-binding) site, located in the N-terminal regulatory domain (Figure 3-9A). These sites are thought to bind DNA independently, and the type of DNA bound at each site dictates the progress of enzymatic activity (74). The N-terminal DNA binding domain is thought to suppress *de novo* methylation, contributing to the preference of the enzyme toward hemi-methylated DNA, and its role as a maintenance methyltransferase (74).

II. Enzymatic activity and mechanism of action for nuclear DNMT1

Methylation of DNA occurs via transfer of a methyl group from a cofactor, *S*-adenosyl-L-methionine (SAM), to the C5 position of a target cytosine residue by one of the DNA (cytosine-5) methyltransferases. The mechanism can be described as having five steps: 1) binding, 2) base-flipping, 3) covalent catalysis, 4) methyl group transfer, and 5) β-elimination (130, 134, 135). The schematic shown in Figure 4-1 starts with Step 3 for simplicity. The first step involves binding of the DNMT to its target CpG site in duplex DNA, which induces base-flipping of the target cytosine out of the DNA helix. During this step, the cytosine is flipped 180° out of the DNA helix, and into the active site pocket of the DNMT (136). The thiol group of the Cys residue within the DNMT ProCys active site forms a covalent thioester bond with the C6 position of the base ring of the target cytosine (134, 135). Methyl group transfer from the cofactor SAM to the C5 position of the target residue occurs by way of an S_N2 nucleophilic substitution mechanism, and is the rate-limiting step of the sequential reaction. Finally, β-

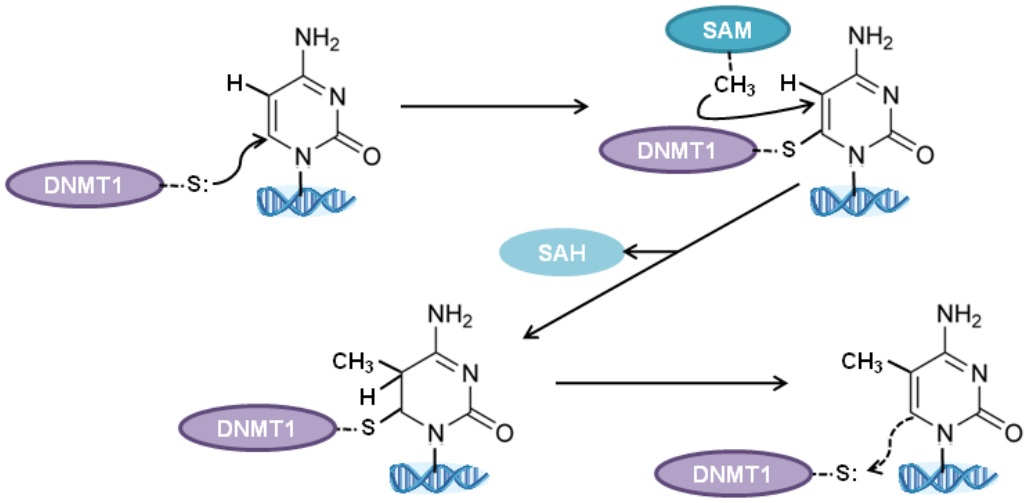


Figure 4-1: Mechanism of DNA methylation. DNA methylation involves covalent complex formation between the target cytosine residue and the sulfhydryl group of the cysteine within the ProCys active site. The methyl group transfer reaction occurs with the aid of methyl group donor, S-adenosyl-L-methionine (SAM) (upper right). DNMT1 is released from the 5mC-modified cytosine residue by β-elimination.

elimination releases the DNMT enzyme from the 5mC-modified base, and S-adenosyl-homocysteine (SAH) is yielded as a by-product of the reaction.

Studies by Bestor and Ingram (136) were the first to suggest that DNMT1 operates in a processive manner. They, and others, have shown that DNMT1 exhibits a higher methylation rate on DNA of greater length, with levels of processivity nearing 99% on hemimethylated DNA substrates (131, 137). DNMT1 remains in intimate contact with the target DNA strand as it slides along, which helps the enzyme maintain its orientation with respect to DNA, as evidenced by its low frequency of skipping target sites (131, 137).

Interestingly, DNMT1 does not behave according to Michaelis-Menten kinetics (118, 132, 138, 139). Rather, the enzyme is capable of more complex kinetics, whereby DNA can act as both a substrate for and a regulator of the reaction (138). This observation implies that the N-terminal part of the enzyme influences the catalytic (C-terminal) domain through allosteric activation, involving two distinct DNA-binding sites, as described above (74, 138, 139). This coordinated interaction between the N-terminus and C-terminus of DNMT1 simultaneously with DNA is supported by evidence that expression of the catalytic domain of DNMT1 alone is incapable of methylating DNA (74). Studies characterizing the allosteric (N-terminal) DNA-binding site uncovered some interesting findings. Deletion of first 580 amino acids within the N-terminal domain has no negative effect on catalytic activity; in fact, expression of an N-terminally truncated protein, lacking the first 501 amino acids, has a 3-10-fold faster catalytic rate than the full-length enzyme against all DNA substrates tested (139). This suggests that the enzymatic activity of DNMT1 is tolerant of limited proteolysis to the N-terminus, which we predict, from data presented in Chapter 2, happens to the mitochondrial isoform of DNMT1 upon translocation into the organelle.

III. Pharmacologic inhibition of total DNMT1 activity

5-aza-cytidine (5-aza-C) and 5-aza-2'-deoxycytidine (5-aza-dC) are nucleoside analogs of cytidine, containing modifications in the 5 position of the pyrimidine ring (Figure 4-2A) (40). Both analogs are potent mechanism-based inhibitors of DNA methylation that function by virtue of their incorporation into DNA (Figure 4-2B). These compounds have been widely used to induce re-expression of genes silenced by methylation and to stimulate cellular differentiation. However, because 5-aza-dC is as effective at one-tenth the dose of 5-aza-C, and because 5-aza-C is considerably more cytotoxic, the deoxy analog is favored as a therapeutic agent (40, 140).

These drugs alter the DNA methylation status of DNA by affecting the expression and activity of the DNA methyltransferases. There are thought to be two modes of action for 5-aza-dC on DNMT1: 1) covalent trapping and 2) proteasomal degradation (141). The first mechanism affects only DNMT1 that is bound by DNA (Figure 4-2B). Modification at the N5 position of an aza-substituted cytidine residue prevents substitution of CH_3 from SAM and β-elimination of the bound DNMT1 enzyme. Since DNMT1 acts processively, covalent trapping of the enzyme results in a dramatic decrease in maintenance methylation activity (141, 142). The second mechanism of action for 5-aza-dC is less understood. The soluble (free) form of DNMT1 is believed to be the primary target for 5-aza-dC-induced proteasomal degradation, although the precise mechanisms underlying degradation are not clear (141).

IV. Objectives

The data presented in Chapter 2 identified a mitochondrial isoform of DNMT1, validated the existence of 5mC modifications within mtDNA, and suggested that expression of mtDNMT1 induced changes in mitochondrial transcription. Using the analogy of the nucleus, we

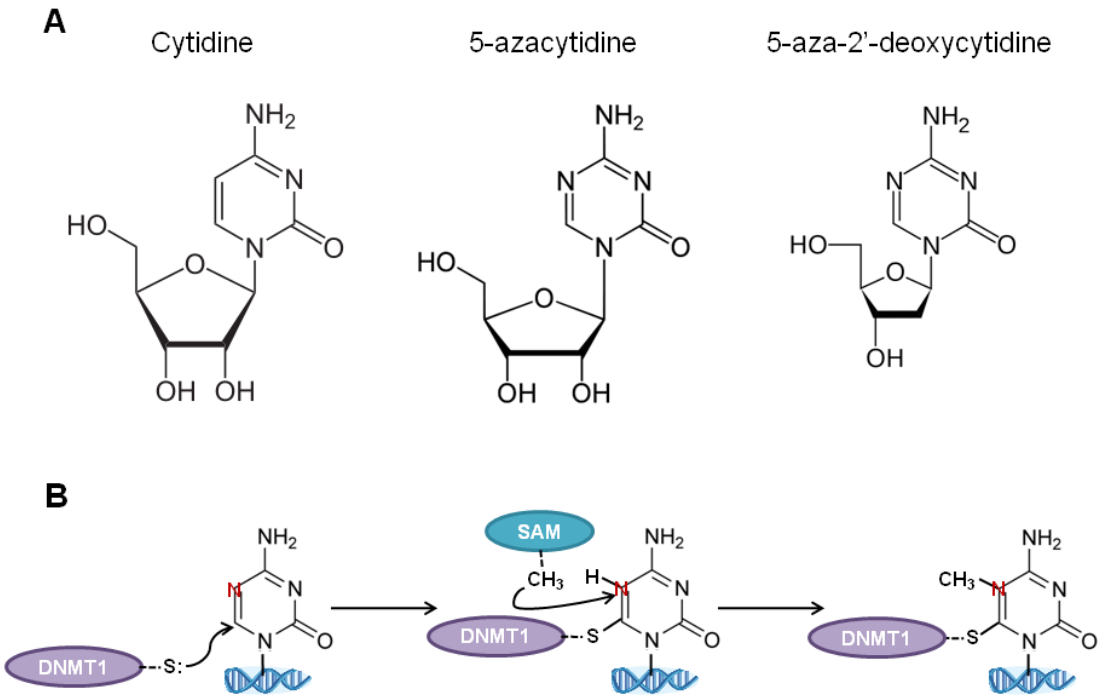


Figure 4-2: Structures of cytidine analogs and mechanism of inhibition of DNMT1 by covalent trapping. (A) Structures for cytidine and nucleoside analogs 5-azacytidine (5-aza-C) and 5-aza-2'-deoxycytidine (5-aza-dC) are shown. (B) Mechanism for DNMT1 inhibition by covalent trapping of the enzyme at 5-aza-C or 5-aza-dC substituted residues.

hypothesized that mtDNMT1 was functioning as a methyltransferase enzyme, modifying actively-replicating mtDNA, and in a role secondary to methylation, also ultimately regulating mitochondrial transcription. Thus, the goal of the studies described below was to provide the missing link between the observations made in Chapter 2, and determine whether mtDNMT1 is catalytically active, and therefore capable of generating the 5mC modification in mtDNA.

MATERIALS AND METHODS

Materials

Cell lines, reagents, antibodies and other materials were obtained as described in Chapter 2. Fluorescently-labeled, internally-modified DNA oligos for the DNMT1 enzyme assay were purchased from Integrated DNA Technologies (IDT, Coralville, IA), and sequences are listed in Table 4-1. *Gla*I enzyme (cat # E494) was purchased from SibEnzyme (West Roxbury, MA). Bacterial CpG methyltransferase *M.Sss*I (cat # M0226L) and recombinant human DNMT1 enzyme (cat # M0230S) were purchased from New England BioLabs (Ipswich, MA). Fluorescence readings in enzyme assays were measured using a BioRad DNA Engine Peltier thermal cycler fitted with a Chromo4 Real-Time Fluorescence Detector attachment. Enzyme assay data was analyzed with Opticon Monitor 3 software.

Methods

Cell culture

HCT116neo^{S} cells and HCT116-TAP cells were grown at 37°C in 5% CO_2 and fed RPMI 1640 (Gibco/Invitrogen) medium, supplemented with 10% FBS, 100U/mL Penicillin and 100mg/mL Streptomycin. HEK293 cells were grown at 37°C in 10% CO_2 and fed DMEM

Table 4-1: Sequences for fluorescently-labeled oligos used in enzyme assay and additional mtIP primers

Species	Primer Set	Forward	Reverse
		Mitochondrial Immunoprecipitation (mtIP)	
Human	ND1	5'-TGCGAGCAGTAGCCAAACAAT-3'	5'-TGATGGCAGGAGTAATCAGAGG-3'
	ND6	5'-AAACACTCACCAAGACCTCAACCC-3'	5'-ATTGATTGTTAGCGGTGTGGTCGG-3'
	No CpG	5'-CTGGTGATAGCTGGTTGTCCAAGA-3'	5'-CCTAGTGTCTAAAGAGCTGTTCCT-3'
		mtDNMT1 Catalytic Activity Assay	
	Oligo1	5'-/6-FAM/CCTATGCG/5mC/ATCAGTTTTCTGATG/5mC/G/5mC/ATAGG/BHQ_1/-3'	
	Oligo2	5'-/6-FAM/CCTATG/5mC/G/5mC/ATCAGTTTTCTGATG/5mC/G/5mC/ATAGG/BHQ_1/-3'	

(Gibco/Invitrogen) medium, supplemented with 10% FBS, 100U/mL Penicillin and 100mg/mL Streptomycin. Cell cultures were fed every other day, and stock cultures were maintained at subconfluent cell densities.

For experimental consistency and because DNMT1 is cell-cycle regulated (78, 92) all cells were treated the same way in the days leading up to an experiment. Cells were fed 48 hrs prior to harvest, split and seeded 24 hrs prior to harvest, and harvested approximately 24 hrs after plating to ensure that the vast majority of cells were still in S-phase, and expressing the maximal level of DNMT1.

Subcellular fractionation and immunoblotting of HCT116-TAP cells

Each subcellular fraction was lysed, quantitated and its constituent proteins were resolved by SDS-PAGE as described in Chapter 2. Gels were transferred to Immobilon PVDF membranes using the wet transfer protocol, as described in Chapter 3. An antibody directed against the far N-terminal epitope (amino acids 1-10) of DNMT1 (Abcam, cat # ab16632) was used for subcellular localization studies because it provided the strongest signal for full-length DNMT1 in the mitochondrial fraction. Antibodies against compartment-specific markers were used to demonstrate the purity of each subcellular fraction: VDAC (Sigma) as a mitochondrial marker and H3K4me3 (Upstate Biotechnology) as a nuclear marker.

Mitochondrial DNA nucleoid extraction

mtDNA nucleoids were isolated from purified mitochondria using an adaptation of the protocol decribed by Garrido, *et al.* (8). Mitochondria were isolated from HCT116-TAP cells as described in Chapter 2. The intact mitochondrial pellet was treated with trypsin (see below)

before being resuspended in 1.5x NE2 buffer (0.25M sucrose, 20mM Tris-HCl, pH 7.5, 2mM EDTA, 7mM β-mercaptoethanol and complete protease inhibitors), and diluted with an equal volume of 0.5X NE2 buffer to a final concentration of 5-7mg/mL of mitochondrial protein. Mitochondria were lysed by adding 20% NP-40 to a final concentration of 0.5%. 1M Spermidine (Sigma, cat # S0266) was added to a final concentration of 3mM to bind and precipitate the mtDNA. Mitochondrial extracts were incubated on ice for 15min to ensure complete lysis, then split evenly between two 1.5mL microcentrifuge tubes: one for immunoblot analysis, the other for enzyme assay. The mitochondrial lysate was sub-fractionated into soluble ("S") and insoluble pellet ("P") fractions by centrifugation at 12,000xg for 20min at 4°C. The soluble ("S") supernatants were divided evenly between two clean 1.5mL microcentrifuge tubes. One aliquot of the "S" fraction was stored as-is at -80°C for immunoblot. The NaCl concentration was adjusted to 200mM and glycerol was added to 50% of the total volume of the other "S" aliquot. 25µL of SDS lysis buffer (62.5 mM Tris pH 6.8, 5% glycerol, 2% SDS, 5% β-mercaptoethanol, and 1x complete protease inhibitor cocktail (Roche)) was added to one aliquot of the "P" fraction for immunoblot analysis. The aliquot of the "P" fraction for enzyme assay was resuspended in 25µL of mtDNMT1 Lysis buffer (50mM Tris-HCl, pH 7.5, 1mM DTT, 1mM EDTA, 5% glycerol and 0.1% Tween 20, 2x complete protease inhibitor cocktail), and the NaCl concentration was adjusted to 200mM and glycerol was added to 50% of the total volume.

Mitochondrial DNA immunoprecipitation (mtIP)

Mitochondrial DNA immunoprecipitation was performed using an adaptation of the protocol described by Lu, *et al.* (143). The ProtA domain of the TAP-tag was employed to isolate mtDNMT1-TAP from crosslinked mitochondrial extracts to ask if and where mtDNMT1

was bound to mtDNA. Mitochondrial extract from non-tagged HCT116 cells was included as a negative control.

Isolation and cross-linking of mitochondria

Mitochondria were isolated from non-tagged HCT116 and HCT116-TAP cells by dounce homogenization and differential centrifugation as described in Chapter 2. Mitochondrial pellets were washed twice with 1mL of mitochondrial homogenization buffer in the presence of complete protease inhibitors (cOmplete EDTA-free Protease Inhibitor Cocktail tablets, Roche, cat # 11873580001), and once with 1mL of mitochondrial homogenization buffer in the absence of protease inhibitors. To washed mitochondrial pellets, Formaldehyde crosslinking solution (11% formaldehyde, 50mM HEPES, pH 8.0, 100mM NaCl, 1mM EDTA, 0.5mM EGTA) was added to $1/10^{th}$ the final volume, and incubated for 15min at 37°C. Cross-linking was terminated by adding glycine to a final concentration of 125mM and incubating at room temperature for 10min. Crosslinked mitochondria were pelleted by centrifugation at 10,000xg for 5min at 4°C, and washed once with 1mL of cold 1x PBS, pH 7.4. Mitochondrial pellets were split into two 1.5mL microcentrifuge tubes and re-pelleted by centrifugation, then lysed in 1mL mtIP Lysis buffer (0.5% Triton X-100, 300mM NaCl, 50mM Tris-HCl, pH 7.4, 100µg/mL leupeptin, 200µM PMSF, complete protease inhibitor cocktail). The mitochondrial lysate was pipeted up and down several times to mix thoroughly, and incubated on ice for 15min with occasional inversion of the tubes, to allow complete lysis.

Sonication of mtDNA

The crosslinked mitochondrial DNA was sonicated using the Diagenode Bioruptor water bath sonicator at the highest pulse, cycling between 30 sec on and 30 sec off, for two rounds 15min each. DNA fragmentation was assessed on a 1% TAE-EtBr agarose gel to verify proper

shearing to 200-700bps. 10µL of sonicated DNA was run alongside 1µg of unsonicated gDNA. The total protein concentration contained within the mitochondrial lysates was determined by Bradford method using a BioRad protein dye (cat # 500-0006), against a standard curve generated by absorbance values measured for known concentrations of BSA (Fisher, cat # 23210) read at 595nm.

Immunoprecipitation of DNMT1-TAP with IgG beads

IgG sepharose beads (GE Healthcare, cat # 17-0969-01) were pipeted using a 2-200µL tip with the very end cut off. 125µL of IgG beads were washed with 1mL mtIP lysis buffer in the absence of complete protease inhibitors three times, spinning at 2000rpm for 2 minutes at 4°C between washes. Equilibrated IgG beads were resuspended into a 50% bead slurry with mtIP lysis buffer in the presence of complete protease inhibitors, and 100µL was aliquoted into two 1.5mL microcentrifuge tubes. 750µg of crosslinked mitochondrial lysate from each cell type was added to the tube containing IgG beads, and incubated at 4°C overnight with end-over-end rotation.

After overnight incubation, IgG beads were spun down at 2000rpm, for 2min at 4°C, and the flowthrough supernatant was collected and retained in a separate 1.5mL microcentrifuge tube labeled “input”. The beads were washed as follows: three times with 1mL of RIPA-140 buffer (10mM Tris-HCl, pH 8.0, 1% Triton X-100, 0.1% SDS, 0.1% deoxycholate, 1mM EDTA, 1mM EGTA, 140mM NaCl); three times with 1mL of RIPA-500 buffer (10mM Tris-HCl, pH 8.0, 1% Triton X-100, 0.1% SDS, 0.1% deoxycholate, 1mM EDTA, 1mM EGTA, 500mM NaCl); three times with 1mL of LiCl buffer (10mM Tris-HCl, pH 8.0, 0.25M LiCl, 0.5% NP-40, 0.5% deoxycholate, 1mM EDTA); and twice with 1mL of TE buffer (10mM Tris-HCl, pH 8.0, 1mM EDTA), pelleting the beads by centrifugation at 2000rpm for 2min at 4°C between each wash.

200µL TE buffer containing 0.5% SDS was added to all tubes and the samples were incubated at 65°C overnight to reverse the crosslinks.

DNA purification

After the reversal of crosslinks, IgG beads were pelleted by centrifugation at 2000rpm for 2min at 4°C. This time, the supernatant contains the immunoprecipitated material, released from the beads after reversing the crosslinks, so the supernatant from each sample was collected and retained. DNA was precipitated by adding $1/10^{th}$ the total volume of 3M NaOAc, 20µg UltraPure glycogen (Invitrogen, cat # 10814010) and 2-3 volumes of 100% ethanol. Tubes were mixed well by vortexing, and stored at -80°C for approximately 2hrs. The DNA was pelleted by centrifugation at 13,000xg for 15min at room temperature. Pelleted material was washed once with 1mL 70% ethanol and the pellets were allowed to air dry on the benchtop for approximately 10-15min. DNA pellets were resuspended in 200µL of HPLC H2O containing 50µg/mL RNase A (Sigma, cat # R5503-100MG) and incubated at 37°C for 1hr. 100µg/mL of Proteinase K (20mg/mL stock, Bioline, cat # BIO-37037) was added to each tube, along with 10% SDS to a final concentration of 0.25%, and samples were vortexed and quickspun to mix thoroughly. Samples were incubated at 37°C overnight to degrade all proteins present.

After overnight incubation, the DNA was subjected to Phenol/Chloroform extraction twice, followed by chloroform/isoamyl alcohol extraction and ethanol precipitation with UltraPure glycogen, as described in Chapter 2. Precipitated DNA pellets were washed once with 1mL of 70% ethanol, and resuspended in 50µL of 10mM Tris, pH 8.0. DNA samples were incubated at 37°C for approximately one hour to help solubilize the pellets.

qPCR with mitochondrial-specific primers

Input DNA samples were quantitated on the NanoDrop spectrophotometer and diluted to make 100µL working stocks of 10ng/µL, 1ng/µL and 0.1ng/µL each to be used as a standard curve. Endpoint PCR was used to confirm that the experiment worked by demonstrating a clear difference in band intensity for products representing a specific interaction with mtDNA over non-specific or negative controls. Endpoint PCR was performed using HotStarTaq Plus Master Mix (Qiagen, cat # 203645), the mtIP3 primer set (Table 4-1) and 1µL of immunoprecipitated DNA from each sample. PCR products were visualized on a 1% TAE-EtBr agarose gel by UV transillumination. Quantitative PCR was performed to quantitate the difference between a specific mtDNMT1-mtDNA interaction over signal from non-specific background. 25µL reactions were prepared using Quantitect SYBR Green PCR Mastermix, 0.6µM of each mitochondrial-specific primer (Table 4-1) and 1µL of immunoprecipitated DNA, and tubes were tapped to mix and quickspun to collect all contents at the bottom of the tube. qPCR was performed using BioRad Peltier thermal cyclers, with PCR cycling conditions specific to each primer set (Table 4-1). Reactions were run in triplicate and reported dilution values averaged to obtain the mean and standard deviation.

Preparation of whole cell and purified nuclear enzyme samples

Approximately $45x10^6$ HCT116-TAP cells were washed twice with ice-cold 1x PBS, pH 7.4 and harvested by scraping. 5% of the whole cell pellet was resuspended in mtDNMT1 Lysis buffer (50mM Tris-HCl, pH 7.5, 1mM DTT, 1mM EDTA, 5% glycerol and 0.1% Tween 20), NaCl was added to 200mM final concentration, and glycerol was added to 50% of the total volume.

Nuclei were isolated from the nuclear/unbroken cell pellet obtained after dounce homogenization during mitochondrial isolation, as described in Chapter 2. The nuclear/unbroken cell pellet was resuspended in 3mL of Buffer S1 (10mM KCl, 10mM HEPES, pH 7.5, 10mM MgCl2, 10% glycerol, 0.25M sucrose, 0.1% Triton X-100, 1mM DTT, 5µg/mL aprotinin, 5µg/mL leupeptin, 5µg/mL pepstatin, 1mM PMSF, 1mM sodium orthovanadate, and 5mM NaF) and overlaid on top of 3mL of Buffer S3 (10mM KCl, 10mM HEPES, pH 7.5, 0.5mM MgCl2, 10% glycerol, 0.88M sucrose, 0.1% Triton X-100, 1mM DTT, 5µg/mL aprotinin, 5µg/mL leupeptin, 5µg/mL pepstatin, 1mM PMSF, 1mM sodium orthovanadate, and 5mM NaF) in a 15mL conical tube. Nuclei were pelleted by centrifugation at 3000rpm in the Beckman J6-MI centrifuge for 15min at 4°C. The nuclear pellet was washed once with 1mL Buffer S3 and divided evenly between two tubes. The nuclear sample in one tube was pelleted, weighed and resuspended in 10x the pellet weight in SDS lysis buffer (62.5 mM Tris pH 6.8, 5% glycerol, 2% SDS, 5% β-mercaptoethanol, and 1x complete protease inhibitor cocktail (Roche)). The other nuclear sample was pelleted, resuspended in mtDNMT1 Lysis buffer, the NaCl concentration was adjusted to 200mM and glycerol was added to 50% of the total volume.

Preparation of mitochondrial DNMT1 enzymes

The storage conditions for purified mitochondrial DNMT1 enzyme preparations were modeled after the recombinant hDNMT1 enzyme (New England BioLabs, cat # M0230S). All mitochondrial DNMT1 enzymes were stored at -20°C as 50% glycerol stocks in 50mM Tris-HCl, pH 7.5, 200mM NaCl, 1mM DTT, 1mM EDTA and 0.1% Tween 20, in the presence of complete protease inhibitors.

Crude mitochondrial extracts

Crude mitochondria were prepared as described in Chapter 2. Mitochondrial pellets were lysed in mtDNMT1 Lysis buffer (50mM Tris-HCl, pH 7.5, 1mM DTT, 1mM EDTA, 5% glycerol and 0.1% Tween 20) in the presence of complete protease inhibitors. Prior to storage, the NaCl concentration was adjusted to 200mM, and glycerol was added to 50% of the total volume. Enzymes were stored at -20°C until analyzed by enzyme assay.

Percoll-purified mitochondrial extracts

Percoll-purified mitochondrial extracts were prepared as described in Chapter 2. These highly-purified mitochondrial pellets were lysed and stored as described above for crude mitochondrial enzymes.

Trypsin treatment of intact mitochondria

Mitochondria were isolated from HCT116 cells by dounce homogenization and differential centrifugation as described in Chapter 2. The crude intact mitochondrial pellet was washed twice with 1mL of mitochondrial homogenization buffer in the presence of complete protease inhibitors (Roche), and pelleted by centrifugation at 10,000xg for 10 minutes at 4°C after each wash. The mitochondrial pellet was washed once with 1mL of mitochondrial homogenization buffer in the absence of protease inhibitors, and resuspended in Trypsin Digestion buffer (10mM HEPES-KOH, pH 7.4, 250mM sucrose, 0.5mM EGTA, 2mM EDTA, 1mM DTT) in a volume 20x the pellet weight. Trypsin-EDTA (Gibco, cat # 25200-056, stock concentration: 100µg/mL, diluted in 1x PBS, pH 7.4), was added to achieve a final concentration of 10µg/mL, and samples were incubated at room temperature for 20 minutes with occasional inversion to mix the tube contents. Bovine trypsin inhibitor (Sigma, cat # T0256, stock concentration: 1mg/mL) was added to a final concentration of 10µg/mL, tubes were inverted

several times to mix, and incubated on ice for approximately 10 minutes. Intact mitochondria were pelleted by centrifugation at 10,000xg for 10 minutes at 4°C, and washed twice more with 1mL of mitochondrial homogenization buffer, this time in the presence of both complete protease inhibitors, and 10µg/mL bovine trypsin inhibitor. Washed mitochondrial pellet was weighed and lysed in a volume of SDS Lysis Buffer 10x the pellet weight.

IgG capture of mtDNMT1-TAP

mtDNMT1-TAP was isolated from crude and/or trypsin-treated mitochondrial extracts using IgG beads to selectively bind the ProtA domain of the TAP-tag. Trypsin-treated mitochondrial pellets were resuspended in 0.5mL of mtDNMT1 Lysis buffer and sonicated using a Diagenode Bioruptor water bath sonicator at the highest pulse for 2.5min, cycling between 5sec on and 5sec off. Sonicated lysate was incubated on ice for 15min to ensure complete lysis.

The $MgCl_2$ concentration was adjusted to 10mM final concentration, which is optimal for the *Age*I restriction endonuclease. 5U of *Age*I enzyme (New England BioLabs, cat # R0552S) was added to the mitochondrial lysate and incubated at 37°C for 20min to digest mtDNA and make the mtDNMT1-TAP enzyme more available for capture by IgG beads.

50µL of IgG beads were pipeted using a 2-200µL tip with the end cut off, and washed three times with 200µL of mtDNMT1 lysis buffer, each time pelleting the beads by centrifugation at 250xg for 2min at 4°C. Washed and equilibrated beads were resuspended into a 50% slurry with 50µL of mtDNMT1 lysis buffer. 50µL of IgG bead slurry was added to the 15mL conical tube containing mitochondrial lysate, and the tube was incubated at 4°C overnight with end-over-end rotation.

After overnight incubation, IgG beads were pelleted by centrifugation at 300xg for 3min at 4°C. 1mL of the unbound flowthrough was collected in a 1.5mL microcentrifuge tube and

stored for analysis. IgG beads containing bound mtDNMT1-TAP were washed three times with 0.5mL mtDNMT1 Lysis buffer, each time incubating for 10min at 4°C with end-over-end rotation, and spinning at 300xg for 3min at 4°C. During the last wash, the lysate was split evenly between two 1.5mL microcentrifuge tubes, one for immunoblot analysis, the other for enzyme assay. The aliquot for immunoblot analysis was spun to pellet the beads, and 25µL of SDS lysis buffer was added. The aliquot for enzyme assay was also spun to collect the beads; 25µL of mtDNMT1 Lysis buffer was added, and the NaCl concentration was adjusted to 200mM and glycerol was added to 50% of the total volume.

TAP-tag purification of mtDNMT1-TAP

HCT116-TAP cells grown at approximately 75% confluence (~$15x10^6$ cells per dish) in sixty 150mm dishes were harvested by scraping. Mitochondria were isolated as described in Chapter 2. Mitochondrial pellets were lysed with 1mL of mtDNMT1 Lysis Buffer, then pooled into a 50mL plastic beaker, and the total volume was adjusted to 15mL with mtDNMT1 Lysis buffer. The mitochondrial lysate was sonicated using a Misonix Sonicator 3000 fitted with the probe attachment, cycling between 1sec on, 5sec off, for a total of 25 cycles. The lysate was incubated on ice for 15min to ensure complete lysis. The 15mL total lysate volume was divided evenly between two 15mL conical tubes (~7.5mL each) for immunoprecipitation.

125µL of IgG beads were pipeted using a 2-200µL tip with the end cut off, and washed three times with 250µL of mtDNMT1 lysis buffer, each time pelleting the beads by centrifugation at 250xg for 2min at 4°C. Washed and equilibrated beads were resuspended into a 50% slurry with 125µL of mtDNMT1 lysis buffer. 100µL of bead slurry was added to each of two 15mL conical tubes containing the mitochondrial lysate, and the tubes were incubated overnight at 4°C with end-over-end rotation.

1x TEV buffer was prepared as described in Chapter 3. Following overnight incubation with IgG beads, the beads were pelleted by centrifugation at 250xg for 2min at 4°C. Approximately 1mL of the unbound flowthrough was collected and retained for analysis. IgG beads were transferred to a 1.5mL microcentrifuge tube and washed three times with 1mL of mtDNMT1 Lysis buffer, each time pelleting the beads at 250xg for 2min at 4°C. The beads were then washed three times with 0.5mL 1x TEV buffer, spinning at 250xg for 2min at 4°C in-between each wash.

Washed IgG beads were resuspended in 300µL of 1x TEV buffer, and 100U (10µL) of AcTEV Protease (Invitrogen, cat # 12575015) was added to the tube. The reaction was incubated at room temperature for 2.5hrs, rotating end-over-end. After 2.5hrs, the reaction was spiked with 50U (5µL) additional AcTEV Protease enzyme. The TEV cleavage reaction was allowed to proceed overnight at 4°C with end-over-end rotation to permit complete cleavage of bound mtDNMT1-TAP from the IgG beads.

After overnight incubation in the presence of AcTEV Protease, the IgG beads were pelleted by centrifugation at 250xg for 2min at 4°C. The supernatant, containing the cleaved mtDNMT1-TAP, was transferred to a clean 1.5mL microcentrifuge tube. The concentrations of EDTA and NaCl within the 1x TEV buffer were adjusted to more closely match the storage conditions for the recombinant hDNMT1 enzyme (1mM EDTA, 200mM NaCl, respectively), and glycerol was added to 50% of the total volume. IgG beads containing residual (uncleaved) bound material were boiled in SDS Lysis buffer for 5min, pelleted by centrifugation, and the supernatant was collected and saved for analysis.

mtDNMT1 catalytic activity assay

Using the fluorescence-based enzymatic assay described by Wood, *et al.* (144) as a model, we developed and optimized a similar assay compatible with the Chromo4 fluorescence detection system fitted on our BioRad Real-Time PCR machines. A schematic diagram of the assay is shown in Figure 4-3.

All enzyme assays were prepared in a total volume of 100µL, consisting of 10mM Tris-HCl, pH 7.5, 5mM $MgCl_2$, 1mM DTT, 5% glycerol, 0.1mg/mL BSA, 25-100mM NaCl and 0-1mM S-adenosylmethionine (SAM). Two stocks of Buffer A were prepared and aliquoted for these experiments: one stock of Buffer A-25, containing 25mM NaCl, and Buffer A-100, which contains 100mM NaCl. Fresh stocks of 1x SEB Buffer (SIBEnzyme, 10mM Tris-HCl, pH 8.5, 5mM $MgCl_2$, 10mM NaCl, 1mM β-mercaptoethanol), 0.8U/µL GlaI in 1x SEB Buffer, 1mM SAM, as well as Oligo1 and Oligo2 (various concentrations) were made before each reaction was assembled.

All reactions were run in duplicate, using low-profile, white PCR strips (BioRad, cat # TLS0851), sealed with flat cap strips (BioRad, cat # TCS0803). Reaction tubes were quickspun using the 96-well plate rotor in the Beckman J2-HC centrifuge to collect all contents at the bottom of the tubes. A BioRad Peltier Thermal Cycler was pre-warmed to 37°C prior to loading reactions into the machine. The length of reactions ranged from 1-6hrs, with fluorescence readings taken every 1min.

Oligo design

Two fluorescently-labeled oligo probes (hemi-methylated Oligo1 and fully-methylated Oligo2) were designed that are identical except for a single methylation site at an internal CpG (Table 4-1). Both oligos were labeled at their 5' end with the fluorescein fluorophore 6-FAM and

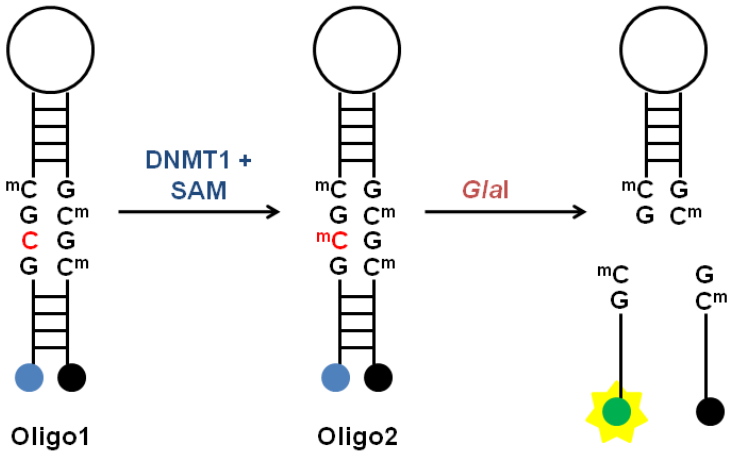


Figure 4-3: Schematic diagram of fluorescence-based assay for methyltransferase activity. Oligos 1 and 2 are fluorescently-labeled at the 5' end with the fluoroscein fluorophore 6-FAM, and at the 3' end with BlackHole Quencher 1 (BHQ1). The oligos form a hairpin, locking the fluorophore moiety in close proximity with the quencher, preventing fluorescence. Within the hairpin structure is a stretch of duplex DNA containing a single methylation site (CpG). When an active methyltransferase is incubated with the hemi-methylated Oligo1, it becomes methylated to form Oligo 2, which is the optimal substrate for *Gla*I restriction endonuclease. Cleavage by *Gla*I releases the fluorophore from the quencher, resulting in an increase in fluorescence that is directly proportional to methylation activity.

at the 3’ end with Black Hole Quencher 1 (BHQ1). Because the excitation/emission wavelengths for the 6-FAM fluorophore mimic that of SYBR Green dye used in qPCRs, the fluorescence could be measured on our real-time PCR machine.

Assay validation and optimal conditions

The fully-methylated Oligo2 was used as a positive control, to determine the optimal concentration of Oligo and GlaI enzyme needed to stay within the linear range of DNMT1 activity. Oligo2 calibration curves were prepared with 0mM SAM, +/- 1U GlaI, Buffer A-100 and a range of Oligo2 concentrations between 0-20nM.

Selectivity of GlaI enzyme between Oligo1 (non-optimal) and Oligo2 (optimal) substrates was measured in duplicate 100µL reactions, with 50µM SAM, 2.4U GlaI, Buffer A-100 and Oligo1 concentrations between 0.028-1µM. An identical range of concentrations was tested for Oligo2, in duplicate 100µL reactions containing 50µM SAM, 0.05U GlaI and Buffer A-100.

Non-specific GlaI cleavage of the hemi-methylated Oligo1 substrate was measured in duplicate 100µL reactions, with 50µM SAM, 0.8U GlaI, +/- 10nM *M.Sss*I, Buffer A-100 and Oligo1 concentrations between 0.13-0.6µM. All validation assays were incubated at 37°C for 1hr, with fluorescence readings taken every 1min.

Methylation activity using recombinant methyltransferases

***M.Sss*I**

Kinetic analysis of *M.Sss*I methylation activity was performed using Oligo1 concentrations ranging from 0.078-0.6µM, in 100µL reactions containing 50µM SAM, 0.8U GlaI, 10nM *M.Sss*I, and Buffer A-100. Reactions were incubated for 2hrs at 37°C, with fluorescence readings taken every 1min.

Recombinant hDNMT1

Kinetic analysis of recombinant hDNMT1 methylation activity was performed using Oligo1 concentrations ranging from 30-240nM, in 100µL reactions containing 1mM SAM, 100µ/mL BSA, 0.8U GlaI, 2.8nM hDNMT1, and Buffer A-25. Reactions were incubated for 2hrs at 37°C, with fluorescence readings taken every 1min.

Methylation activity of mitochondrial DNMT1 enzymes

Mitochondrial enzyme preparations were quantitated by Bradford assay, against a standard curve of known concentrations of BSA read spectrophotometrically at 595nm. Between 0.5-2µg of total mitochondrial lysate was used as the source of mtDNMT1 enzyme when lysates could be quantitated. When protein concentrations were too low to be quantitated accurately, a fixed volume (5-10µL) of mitochondrial enzyme was added to each reaction.

All mitochondrial DNMT1 enzyme preparations were assayed in duplicate 100µL reactions containing the following: 1mM SAM, 0.8U GlaI, 240nM Oligo1 and Buffer A-25. Reactions were incubated at 37°C for 3-6hrs, with fluorescence readings taken every 1min.

Manipulation of mtDNMT1 activity and mtDNA methylation

Treatment with 5-aza-2'-deoxycytidine (5-aza-dC)

HCT116-TAP cells were grown at approximately 50-60% confluence in two 150mm dishes for each condition, per timepoint. Cells were treated with 3µM 5-aza-dC 24hrs after plating, and harvested at 24, 48 and 72hrs after the first 5-aza-dC treatment. Fresh growth medium (RPMI 1640 supplemented with 10% FBS, 100U/mL Penicillin and 100mg/mL Streptomycin) containing fresh drug (3µM 5-aza-dC) was replenished each day. 5-aza-dC treatment was ceased after 48hrs. Cells were harvested from both 150mm dishes per condition at

each timepoint and pooled. Mitochondria were isolated as described in Chapter 2, mitochondrial pellets were further subjected to trypsin treatment and mtDNA nucleoid extraction, as described above. Sub-mitochondrial fractions were divided evenly between two 1.5mL microcentrifuge tubes: one half for analysis by immunoblot, the other half for analysis in an enzymatic assay.

SDS-PAGE and immunoblotting of sub-fractionated mitochondrial extracts

Protein concentrations of "S" and "P" fractions collected from 5-aza-dC-treated mitochondria were determined by Bradford assay, but because the total protein concentrations in these samples was so low, the maximal volume (25µL) of each sample was loaded onto 4-15% gradient SDS-PAGE gels for analysis by immunoblot, as described previously.

Methylation activity of sub-fractionated mitochondrial extracts

Again, because the total protein concentrations for "S" and "P" enzyme preparations for were so low, a fixed volume (5-10µL) of each sample was used as the enzyme source in mtDNMT1 catalytic activity assays, as described above.

"S" and "P" enzyme preparations from 5-aza-dC-treated mitochondria were kept in a small enough volume that they could be quantitated by Bradford assay, so a fixed 2.0µg of each fraction was added to those enzyme activity assays (Figure 4-12 B, C, D).

RESULTS

GlaI activity against Oligo1 and Oligo2

A fluorescently-labeled DNA oligo (Oligo1, Figure 4-3) was designed lacking methylation at a single CpG site, such that a single methylation event by an active methyltransferase would result in the separation of fluorophore and and quencher, and could be directly measured as an increase in fluorescence. A second oligo (Oligo2), identical in sequence,

but with all four CpG sites pre-modified by 5mC, contains the optimal restriction site for *Gla*I (5'-GC^{M}GC^{M}-3'). The tetra-methylated Oligo2 allowed us to optimize the conditions of the assay to work with our fluorescence detection method prior to testing any enzyme samples.

To first assess whether the assay was working in our hands, a calibration curve was run to measure *Gla*I cleavage across increasing concentrations of the fully-methylated Oligo2 substrate. No DNA methyltransferase was needed in these reactions. Endpoint fluorescence values, representing *Gla*I cleavage of Oligo2, were measured across a range of Oligo2 concentrations (6.33-20nM). Fluorescence units per unit of *Gla*I (AU/U) were fitted to a linear function, and the goodness of fit was calculated at $R^2 = 0.9685$ (Figure 4-4A).

The selectivity of *Gla*I between sequences containing either fully-methylated (Oligo2) or hemi-methylated (Oligo1) target restriction sites was measured across various concentrations of either substrate (Figure 4-4B). No methyltransferase was added to these reactions, as we intended to assess just the activity of *Gla*I against each of the DNA substrates. Reported fluorescence values over time were used to calculate the rate of *Gla*I cleavage against either Oligo1 (using a maximal 2.4U *Gla*I per reaction), or Oligo2 (using only 0.05U *Gla*I per reaction). To obtain a stringent assessment of how *Gla*I behaves with both oligo substrates, these reactions were biased in favor of non-specific cleavage by adding nearly 50 times more *Gla*I to the reactions with the hemi-methylated Oligo1. This gave us confidence that the levels of unwanted Oligo1 cleavage we would be detecting in later mitochondrial enzyme assays using only 0.8U *Gla*I per reaction was minimal, and would not interfere with our activity measurements. The initial rates for both sets of reactions were graphed along oligo concentration (Figure 4-4B), and show that with 0.05U *Gla*I the reaction is only linear across 0.028-0.13µM concentrations of Oligo2, and quickly becomes saturated by substrate at Oligo2 concentrations

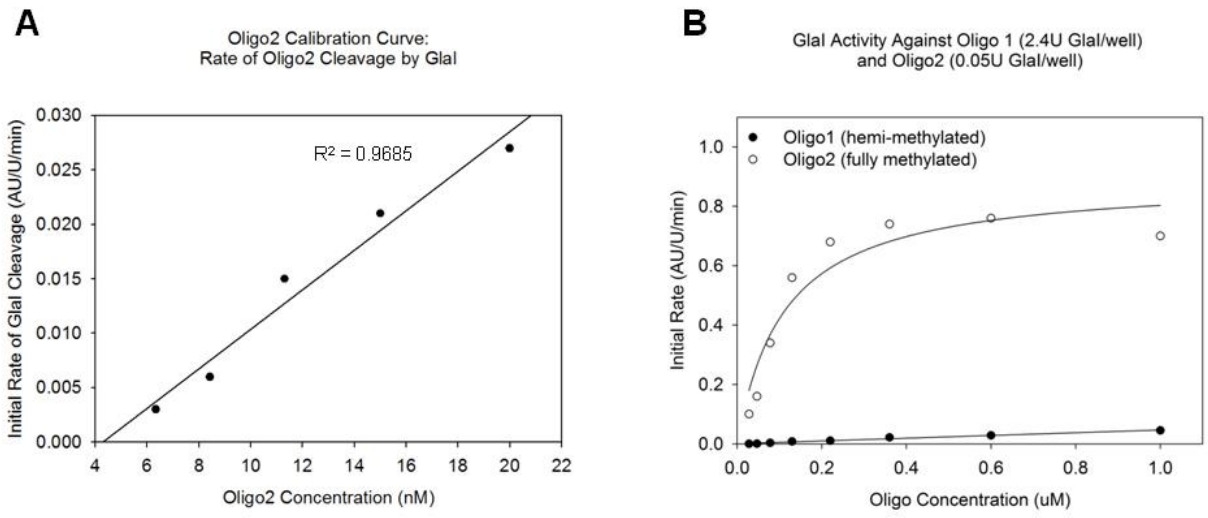


Figure 4-4: Validation of fluorescence-based enzymatic assay and quantitation of non-specific GlaI cleavage. (A) A calibration curve with increasing amounts of the fully-methylated Oligo2 was used to examine the efficiency of *Gla*I cleavage. Initial methylation rates (AU/U *Gla*I/min) were graphed over substrate concentration, and demonstrate strong linearity with R^2 = 0.9685. (B) *Gla*I activity against Oligo1 and Oligo2 were compared. Initial rates of methylation (AU/U *Gla*I/min) were plotted across substrate concentration. *Gla*I cleavage of Oligo2 was as much as 26-fold higher than cleavage of Oligo1.

above that. *Gla*I activity against the non-optimal Oligo1 substrate is linear throughout the entire range of concentrations, however, the rate is well below (by as much as 26-fold) the level of *Gla*I activity against Oligo 2.

Wood, *et al.* (144) determined 0.8U *Gla*I per 100µL reaction to be the optimal concentration of *Gla*I in standard enzyme assays, so this is the amount of *Gla*I used in all of our analyses, unless otherwise mentioned. They also determined that 240nM Oligo1 was the optimal substrate concentration for the kinetic analysis of the commercially-available recombinant DNMT1 enzyme (144), so this is the amount of Oligo1 used in all of our analyses, unless otherwise mentioned.

To get a more accurate measurement for the level of background signal we could expect using the conditions chosen for our mitochondrial DNMT enzyme assays, we measured *Gla*I activity across a range of Oligo1 concentrations (0.13-0.6µM) in both the presence and absence of an active methyltransferase enzyme (*M.Sss*I). The rates were calculated for both reactions, and are shown in Table 4-2. The levels of *Gla*I activity generated by non-specific cleavage of the hemi-methylated Oligo1 in the absence of *M.Sss*I accounted for less than 10% of the total assay signal achieved with *M.Sss*I (Table 4-2), and was less than levels reported by Wood, *et al.* (144). Together, this data demonstrates that *Gla*I exhibits a strong preference for a fully-methylated DNA substrate, and non-specific cleavage of Oligo1 should not impact the quality of our enzyme assay results.

Table 4-2: Non-specific cleavage of the hemi-methylated Oligo1 by GlaI enzyme

	Methylation Rate (AU/U GlaI/min)		
Concentration of Oligo1 (μM)	Positive Control (+M.SssI)	Negative Control (-M.SssI)	% Background Cleavage
0.13	0.0006	-0.0006	-100.00
0.22	0.0069	0.0006	9.09
0.36	0.0075	0.0000	0.00
0.6	0.0256	0.0025	9.76

Methylation activity of *M.Sss*I and recombinant hDNMT1 enzymes is linear across substrate concentration and time

The activities of a two well-characterized CpG methyltransferases were measured using the fluorescence-based enzymatic assay across a range of Oligo1 concentrations. Plotting the steady-state rates for both *M.Sss*I and recombinant hDNMT1 against the concentration of Oligo1 shows a proportional linear increase in rate for both enzymes. The dependence of recombinant hDNMT1 activity on the concentration of Oligo1 substrate is shown in Figure 4-5A, and demonstrates a strong fit to a linear function (R^2 = 0.9918). Similarly, the effect of Oligo1 concentration on the activity of *M.Sss*I is shown in Figure 4-5B, and also demonstrates strong linearity (R^2 = 0.9995). In addition, Figure 4-5C shows that methylation by *M.Sss*I is linear over time. The observations that both of these enzymes exhibit a linear increase in the rate of methylation as a function of substrate concentration is important because it indicates that we have neither saturated nor limited either enzyme (*Gla*I or DNMT1/*M.Sss*I), allowing the reactions to proceed maximally.

Furthermore, as these are coupled reactions, measuring the rates of two separate enzymes (DNMT and *Gla*I), it is important to note that we observed no "lag phase" in fluorescence prior to reaching the linear steady-state of reaction, as Wood, *et al.* (144) described. This indicates that *Gla*I cleavage of the oligo substrate is immediate, and that the reported fluorescence values can be taken to represent true steady-state methylation rates, not rates of *Gla*I activity.

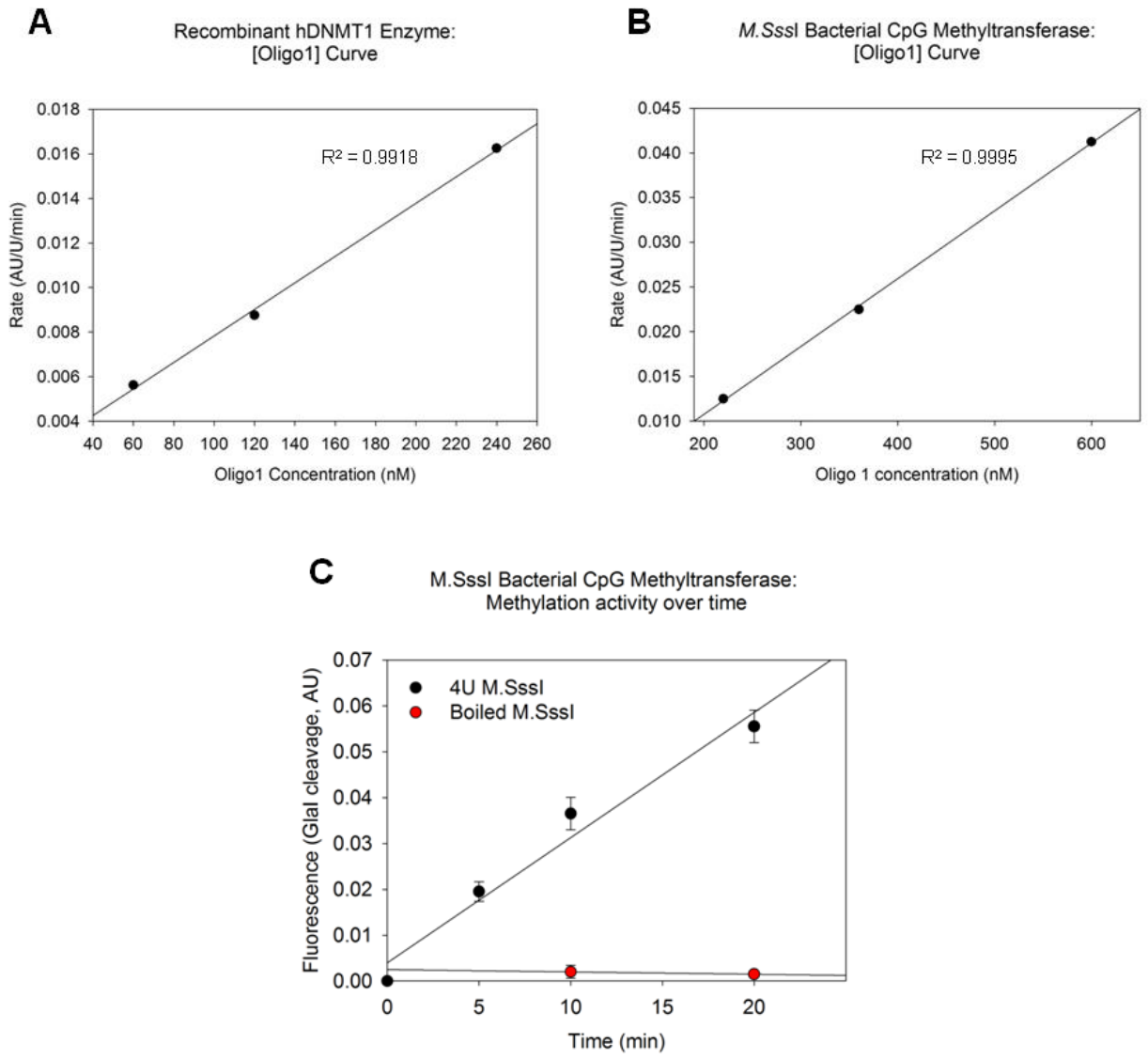


Figure 4-5: Kinetic analyses of *M.Sss*I and recombinant hDNMT1 positive control enzymes. (A) Methylation activity for the recombinant human DNMT1 enzyme is also strongly linear over a range of Oligo1 concentrations, with R^2 = 0.9918. (B) Methylation activity for the bacterial *M.Sss*I methyltransferase was measured relative to Oligo1 concentration, and is strongly linear with R^2 = 0.9995. (C) Methylation activity for 4U of the bacterial *M.Sss*I enzyme was plotted over time and also exhibits linearity.

Methyltransferase activity from whole cells and purified nuclei is linear over time

The methyltransferase activity of a 1:100 dilution of whole cell extract was measured and exhibits a linear increase in *Gla*I cleavage over time (Figure 4-6A). This indicated that the storage and assay conditions we had chosen worked well, not only for purified enzymes purchased from New England BioLabs, but also for enzymes that we prepared in-house. To get an estimate for the level of DNA methyltransferase activity found in purified nuclei, and to approximate the levels we might expect from purified mitochondrial extracts based on previous calculations that mtDNMT1 only represents 1-2% of total DNMT1 in the cell, we measured dilutions of our nuclear enzyme preparations. Purified nuclei demonstrated methyltransferase activity that was linear over time, and proportional to the amount of total nuclear extract assayed (Figure 4-6B).

mtDNMT1 localizes to the insoluble, mtDNA-containing sub-mitochondrial pellet

To observe where mtDNMT1 partitions within the mitochondrial compartment, trypsin-treated mitochondria were sub-fractionated into soluble ("S") and insoluble pellet ("P") fractions using a protocol adapted from Garrido, *et al.* (8). An aliquot (75µg) of whole cell lysate and 18µg of each: mitochondrial, S and P sub-fractions were resolved on a 4-15% gradient gel by SDS-PAGE (Figure 4-7A). The membrane was probed with control antibodies to VDAC and H3K4me3 to demonstrate the purity of each fraction, and with an antibody to DNMT1 to track its sub-mitochondrial localization. At first glance, it is clear that full-length DNMT1 is detected in total mitochondrial extract (M), and both the soluble (S) and insoluble (P) fractions. However, we examined the relative proportion of mtDNMT1 present in either sub-mitochondrial fraction semi-quantitatively, with the hope of gaining insight into how this protein exists in mitochondrial

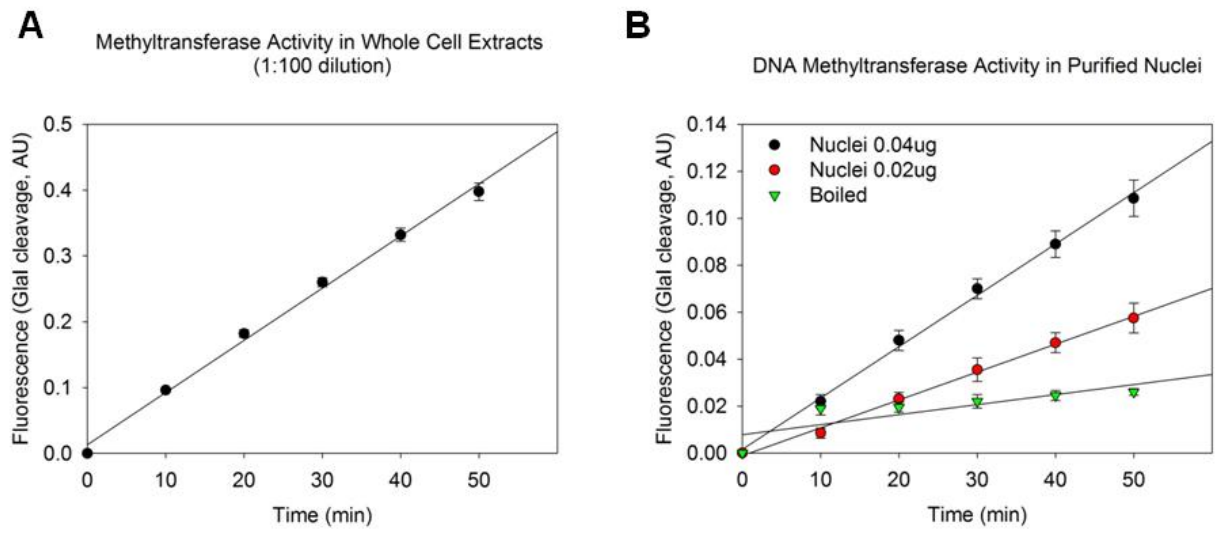


Figure 4-6: Methyltransferase activity assays with whole cell and purified nuclear extracts. Enzyme preparations from whole cells (A) and purified nuclei (B) were diluted to approximate levels of mitochondrial DNMT activity, based on previous estimates that mtDNMT1 represents only 1-2% of the total DNMT1 within a cell. Using standard assay conditions (0.8U/well *Gla*I and 240nM Oligo1), both whole cell and purified nuclear enzymes demonstrate linear methyltransferase activity as a function of time.

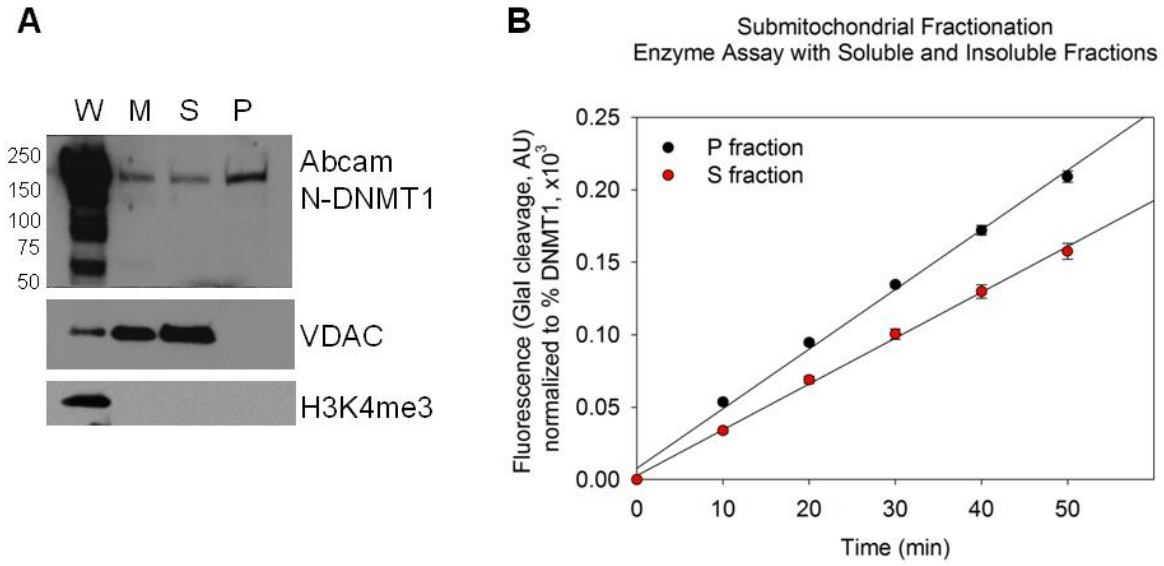


Figure 4-7: mtDNMT1 partitions with the insoluble mtDNA-containing pellet. (A) Mitochondrial DNA nucleoids were isolated as described by Garrido, *et al.* (21) to determine where mtDNMT1 partitions on a sub-mitochondrial level. The soluble (free) supernatant was separated from the insoluble pellet, which contains mtDNA and mtDNA binding proteins. Using an antibody to DNMT1, mtDNMT1 was seen to localize with both soluble "S" and insoluble "P" fractions. Densitometric readings of the immunoblot with protein concentrations determined by Bradford assay allowed calculation of the relative proportion of mtDNMT1 in each sub-mitochondrial fraction: 89.3% of mtDNMT1 is soluble; 10.7% of mtDNMT1 is insoluble. (B) Both soluble and insoluble forms of mtDNMT1 exhibit catalytic activity *in vitro*. Values represent fluorescence readings normalized to the relative proportion of mtDNMT1 in each fraction, as determined by densitometry of the immunoblot in (A).

space, and what levels of methylation activity we might expect from an enzyme that is predicted to be bound to mtDNA. In Figure 4-7A, the protein loaded in the first two lanes represents 75µg of whole cell lysate (W) and 18µg of trypsin-treated mitochondrial lysate (M), respectively, which is standard for all of our subcellular fractionation immunoblots. The S and P lanes were also loaded with 18µg of the total protein collected for either fraction, but this amount of total protein is disproportionate between the two fractions, and this impacts our interpretation of the relative levels of soluble versus insoluble mtDNMT1. The protein loaded in lane "S" represents only 5% of the total S fraction collected, and only 0.33% of the total mitochondrial protein. However, lane "P" represents 51% of the entire insoluble pellet, and 0.91% of the total mitochondrial protein. Using densitometric readings of the signals detected for DNMT1 in either lane, the relative amount of mtDNMT1 in the soluble fraction was calculated to be 89.3%, whereas only 10.7% of mtDNMT1 partitioned with mtDNA. So, despite a more intense signal for mtDNMT1 in the P fraction, the vast majority of mtDNMT1 was soluble and freely available to bind mtDNA, or participate in some other, unknown function. Enzyme assays were run using S and P samples, and as shown in Figure 4-7B, enzymes extracted from both soluble and insoluble mitochondrial pools contain active methyltransferase. Fluorescence readings were normalized to the relative proportion of mtDNMT1 within each sub-mitochondrial fraction, and show that the form of mtDNMT1 purified bound to mtDNA exhibits slightly greater activity than the soluble form.

mtDNMT1 binds mtDNA in a manner that is proportional to CpG density

Data presented in Chapter 3 clearly demonstrated that the DNMT1-TAP allele translocates properly to the mitochondrial compartment. Data presented in Figure 4-6A showed

that mtDNMT1 co-localizes with the insoluble mitochondrial pellet containing mtDNA. We therefore asked whether mtDNMT1 binds to mtDNA in order to understand the function of mtDNMT1. The HCT116-TAP cell line was an ideal tool for asking whether mtDNMT1 interacted with mtDNA.

Using an adaptation of a mitochondrial DNA immunoprecipitation protocol described by Lu, *et al.* (143), we exploited the DNMT1-TAP allele and employed IgG beads to selectively bind the ProtA domain of the TAP-tag. Formaldehyde-crosslinked mitochondrial lysates were immunoprecipitated with IgG beads and qPCR with primers specific for mtDNA (Table 2-2) was used to quantitate the interaction between mtDNMT1 and mtDNA. Immunoprecipitates from HCT116-TAP cells were substantially enriched for mtDNA in comparison with immunoprecipitates from non-tagged parental HCT116 cells, except for an amplicon containing no CpG dinucleotides, which gave an equally low signal from both cell lines (Figure 4-8). These data suggest a CpG-dependent interaction between mtDNMT1 and the mitochondrial genome, and confirm the localization of this protein to the mitochondrial matrix. The mtDNMT1-mtDNA interaction was evident in the D-loop control region, which carries the mitochondrial origin of replication and promoters, as well as in several rRNA and protein coding regions. The level of enrichment was dependent on the target amplicon; five of the six regions probed showed a three- to five-fold enrichment of mtDNA sequences in immunoprecipitated material from HCT116-TAP cells, compared to control. However, qPCR of the region spanning the junction between 12S and 16S rRNA genes (primer set 2) showed only a two-fold enrichment in binding of mtDNMT1-TAP. Interestingly the density of CpG dinucleotides within this amplicon is less than 50% of that in all other amplicons analyzed, suggesting that the interaction of mtDNMT1 with mtDNA is proportional to CpG density.

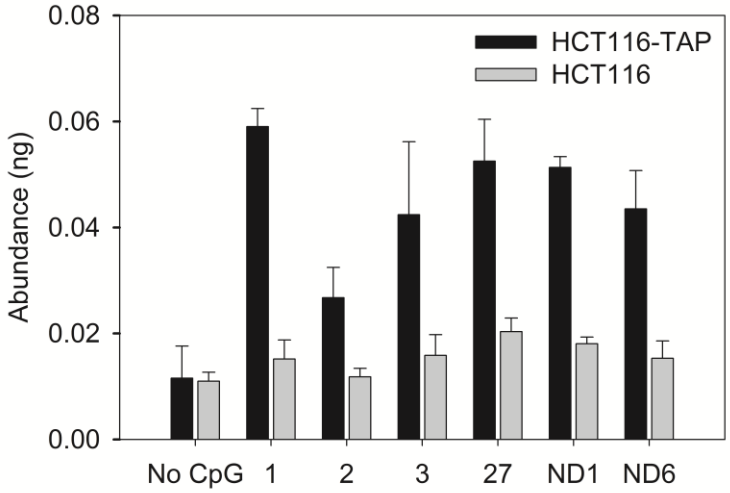


Figure 4-8: mtDNMT1 specifically binds mtDNA in a manner that is proportional to CpG density. mtDNMT1-TAP interacts specifically with various regions along mtDNA. Nonspecific background was assessed using mitochondrial lysates from non-tagged HCT116 cells. Enrichment of mtDNA in immunoprecipitates from each cell type was measured against a standard curve of purified mtDNA, and represent absolute values. Bars represent the mean +/- SD of six replicate samples from two independent experiments.

Highly purified mitochondrial extracts exhibit CpG-specific methyltransferase activity *in vitro*

Mitochondrial extracts, both crude and Percoll-purified preparations, as well as the TAP-tag purified mtDNMT1 enzyme were assayed for methyltransferase activity using 240nM Oligo1. Enzyme activity in mitochondrial lysates (0.5µg) was linear with time for the duration of the assay, as measured by *Gla*I cleavage in fluorescence units (AU) (Figure 4-9A). Linearity of the reactions over time indicates that the reactions progressed maximally, without limitation by substrate concentration. The curves for both crude and Percoll-purified mitochondrial extracts are not remarkably different, suggesting a high degree of purity for our "crude" preparations. Because Percoll purification is regarded as the "gold standard" for isolation of pure mitochondria, the observation that DNA methyltransferase activity in Percoll-purified mitochondrial lysates is not substantially different from crude mitochondrial DNMT activity indicates that the enzyme preparations are similar in purity. The TAP-tag purified mtDNMT1 enzyme also exhibited a linear increase in methylation activity over time; however, the rate for this reaction (0.0083 AU/U *Gla*I/min) was roughly half that for Percoll-purified total mitochondrial lysate (0.0194 AU/U *Gla*I/min). All enzyme preparations demonstrated methyltransferase activity that was also linear as a function of enzyme concentration (Figure 4-9B, C).

mtDNMT1 activity in intact mitochondria is resistant to trypsin treatment

To address concerns that the methylation activity measured from our purified mitochondrial extracts might be nuclear in origin, the intact mitochondrial pellet isolated from non-tagged HCT116 cells was subjected to trypsin digestion, or a mock treatment as described

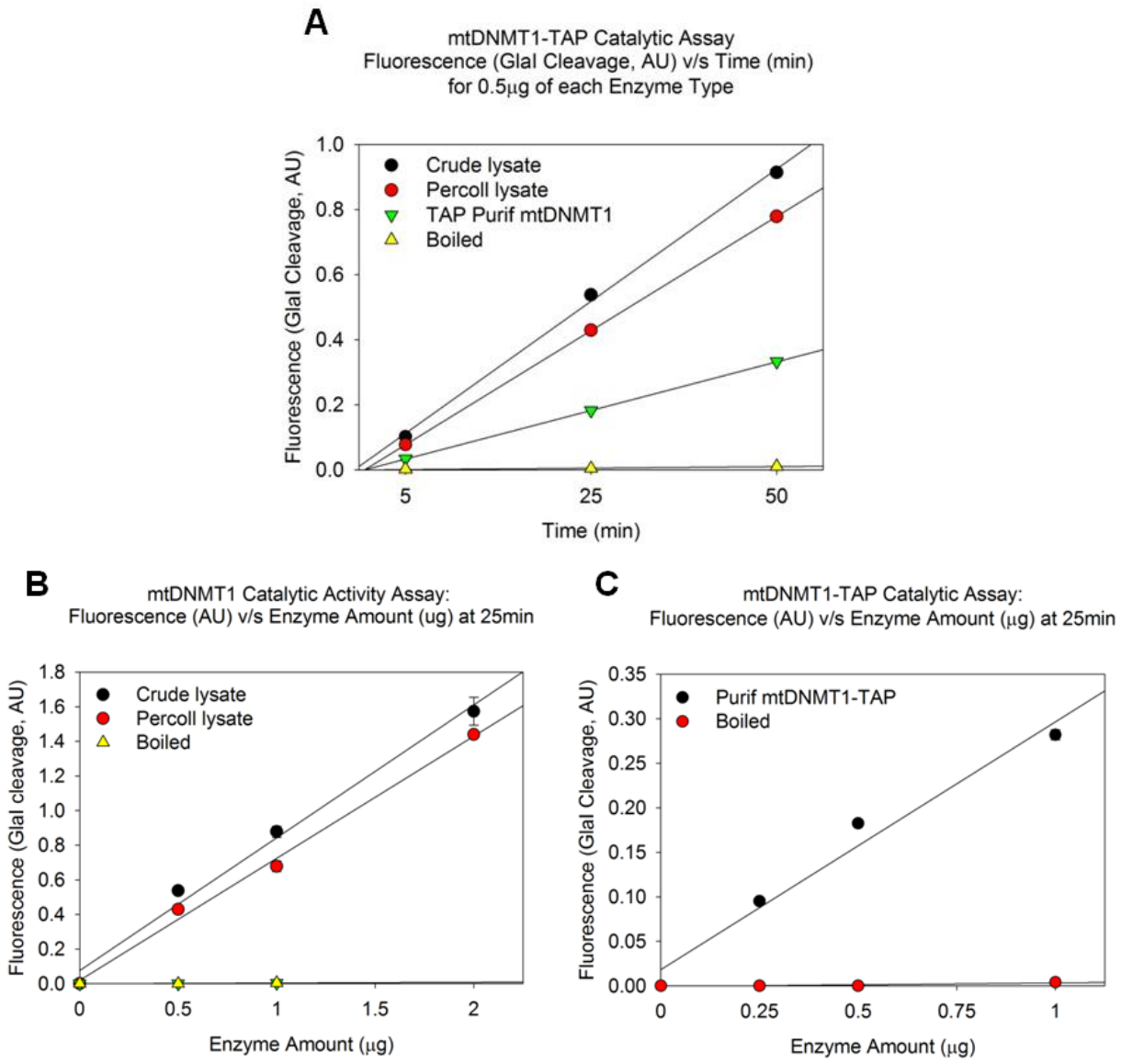


Figure 4-9: Enzyme assays detect CpG-specific methyltransferase activity in highly-purified mitochondrial extracts. (A) 0.5µg of crude or Percoll-purified mitochondrial extracts, along with 0.5µg of TAP-tag purified mtDNMT1-TAP enzyme were assayed for methylation activity using the coupled fluorescence-based enzymatic assay shown in Figure 4-3. Fluorescence values (AU) were graphed relative to time, are linear for the duration of the assay, and represent substantial methylation activity over the boiled enzyme negative control. (B, C) Catalytic activity was also linear with enzyme concentration. Values for crude and Percoll-purified enzymes were separated from the TAP-tag purified mtDNMT1-TAP enzyme for clarity.

by Szczesny, *et al.* (145) prior to lysis. Whole cell lysate (75µg) and 18µg of each mitochondrial lysate were resolved on a 4-15% gradient gel by SDS-PAGE. The membrane was probed with an antibody to VDAC as a loading control, and an H3K4me3 antibody to demonstrate the purity of mitochondrial fractions (Figure 4-10A). An N-terminal DNMT1 antibody showed that, although signal decreased in mock and trypsin-treated mitochondria, DNMT1 could be still easily detected. The continued presence of DNMT1 in mitochondrial extracts following trypsin treatment confirmed its presence within the mitochondrial compartment.

Enzyme preparations were made from the trypsin-treated mitochondrial extracts, and methylation activity was measured for equivalent total protein from each sample relative to time. While there were only subtle differences in activity from untreated, mock and trypsin-treated mitochondria, a sample of mitochondrial extract that was digested with trypsin after lysis showed only background levels of activity (Figure 4-10B).

Mitochondrial extracts from non-tagged and TAP-tagged HCT116 cells demonstrate nearly identical levels of methyltransferase activity

Given the data from Chapter 3, suggesting that total DNMT1-TAP is overexpressed in the HCT116-TAP cells, we wanted to verify that we were not misrepresenting the level of methylation activity in mitochondria by using the HCT116-TAP cell line. Thus, we directly compared the activities of various amounts of either non-tagged or TAP-tagged mitochondrial extracts over time. As shown in Figure 4-11, there is very little difference in the methylation activity of mitochondria from non-tagged and TAP-tagged HCT116 cells. This data also suggests that the TAP-tag is not interfering with activity of the mtDNMT1-TAP allele.

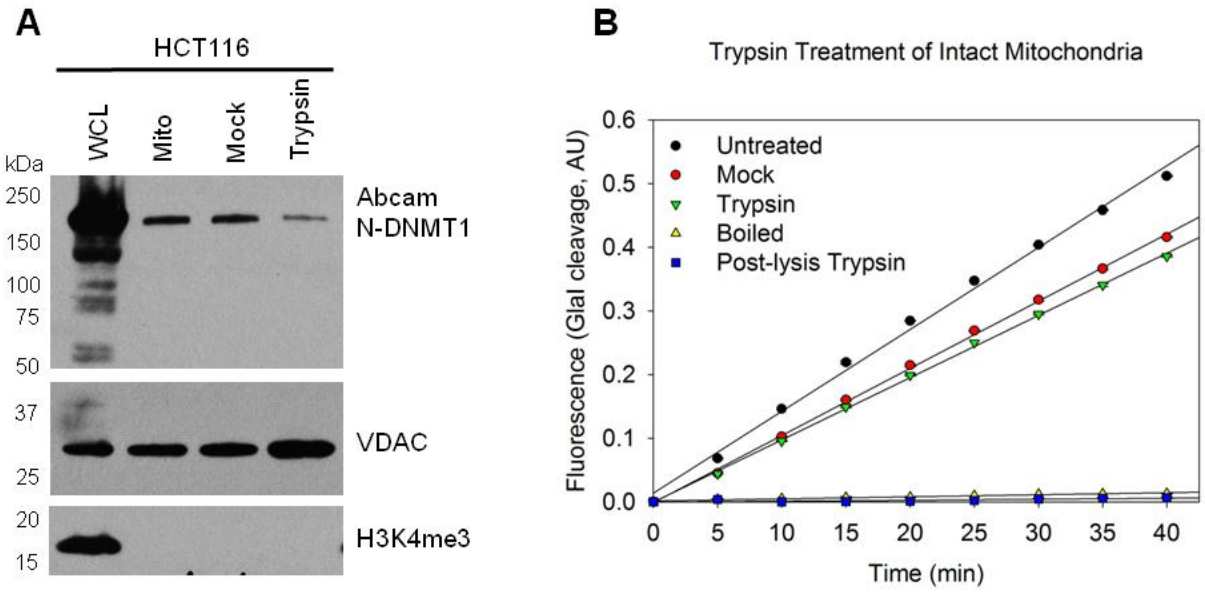


Figure 4-10: Trypsin treatment of intact mitochondria does not affect mtDNMT1 enzyme activity. (A) Immunoblots show the continued presence of full-length DNMT1 in mitochondria subjected to trypsin digestion. VDAC was used as a loading control; H3K4me3 demonstrates purity from contaminating nuclear material. (B) Methylation activity from 2µg of each enzyme preparation was graphed over time, and show only minimal differences in the activity from untreated, mock and trypsin-treated mitochondria.

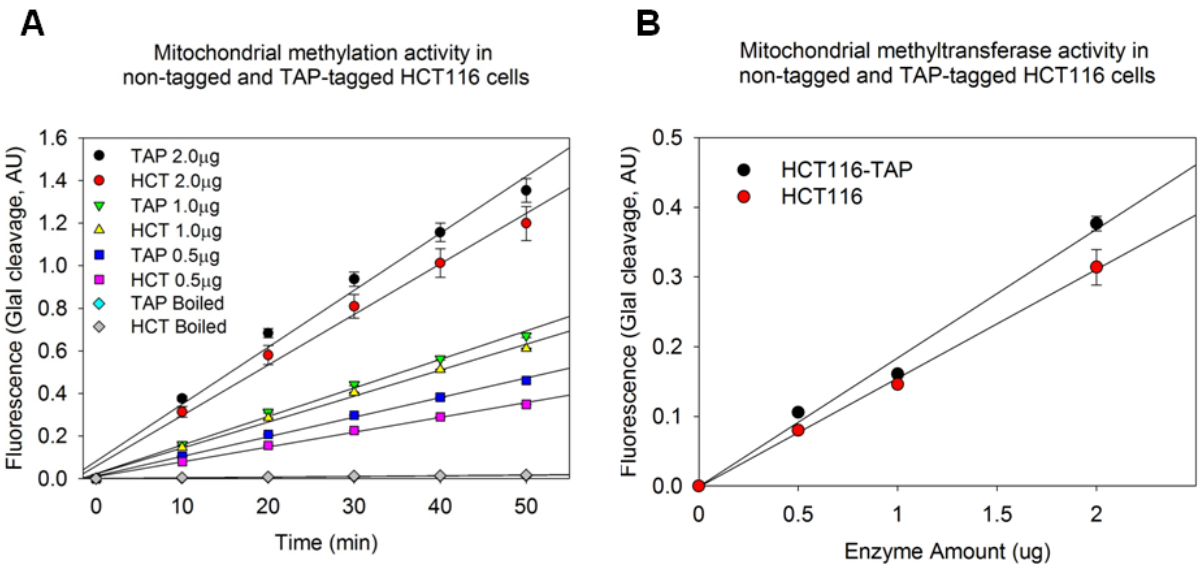


Figure 4-11: Non-tagged and TAP-tagged HCT116s exhibit nearly identical mitochondrial methyltransferase activity. Mitochondria from non-tagged and TAP-tagged HCT116 cells were assayed for methylation activity. (A) Fluorescence values (AU) were graphed over time (A) or enzyme amount (B) at 30min. No substantial differences in catalytic activity were observed for different concentrations of mitochondrial lysate from either cell type.

IgG Capture is highly specific for the mtDNMT1-TAP allele

The specificity of the TAP-tag ProtA-IgG bead interaction was assessed by comparing the enzyme activity within immunoprecipitated material from either non-tagged or TAP-tagged HCT116 mitochondrial lysates, incubated with IgG beads. The level of nonspecific methylation seen in immunoprecipitates from the non-tagged HCT116 cells is barely above the Boiled enzyme negative control, while that for immunoprecipitated material from the TAP-tagged cells is substantially greater (Figure 4-12). This suggests that IgG beads both specifically and selectively bind the TAP-tag, allowing isolation of a functional mtDNMT1-TAP allele.

Restriction digest of mtDNA allows for capture of more mtDNMT1 protein without a corresponding increase in mtDNMT1 activity

Trypsin-treated mitochondrial extracts were incubated with *Age*I restriction endonuclease to digest mtDNA that might be bound to mtDNMT1; a property we hypothesized was preventing efficient capture and enzymatic analysis of the mtDNMT1 enzyme in previous experiments (Figure 4-9). *Age*I was chosen for the relatively high frequency of recognition sites within the mitochondrial genome. Because the protein concentration for the *Age*I IgG-captured sample was below the limits of detection by Bradford assay, the maximum volume (25µL) of that sample was loaded onto a 4-15% SDS-PAGE gel, alongside 18µg of either untreated or trypsin-treated mitochondrial lysate. Immunoblot analysis using an antibody to DNMT1 shows a strong signal for mtDNMT1 in all three mitochondrial extracts (Figure 4-13A). The intensity of the protein band visible in lane 3, representing mtDNMT1-TAP captured on IgG beads after *Age*I digest of mtDNA, is substantially stronger than we have obtained using the IgG capture technique alone, without its prior digestion. This suggests that mtDNMT1-TAP exists tightly bound to mtDNA,

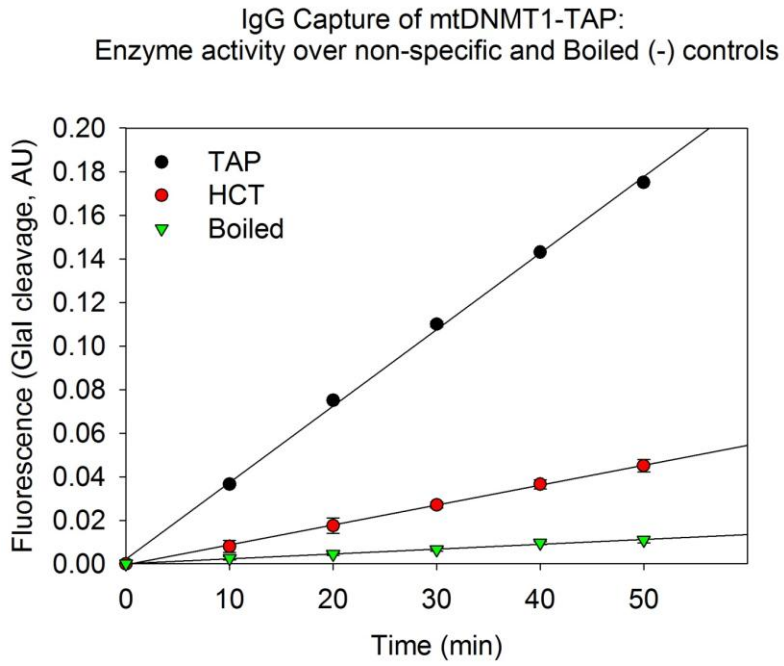


Figure 4-12: Specificity of IgG capture for the mtDNMT1-TAP allele. Mitochondria were isolated from non-tagged and TAP-tagged HCT116 cells and immunoprecipitated with IgG sepharose beads, which selectively bind the ProtA domain of the TAP-tag. A boiled enzyme negative control measured background fluorescence. Non-specific binding between IgG beads and mitochondrial lysate from non-tagged HCT116 cells showed only minimal levels of methylation activity. The substantially greater methylation activity in immunoprecipitates from HCT116-TAP cells demonstrates that IgG beads are selective for isolating the TAP-tag.

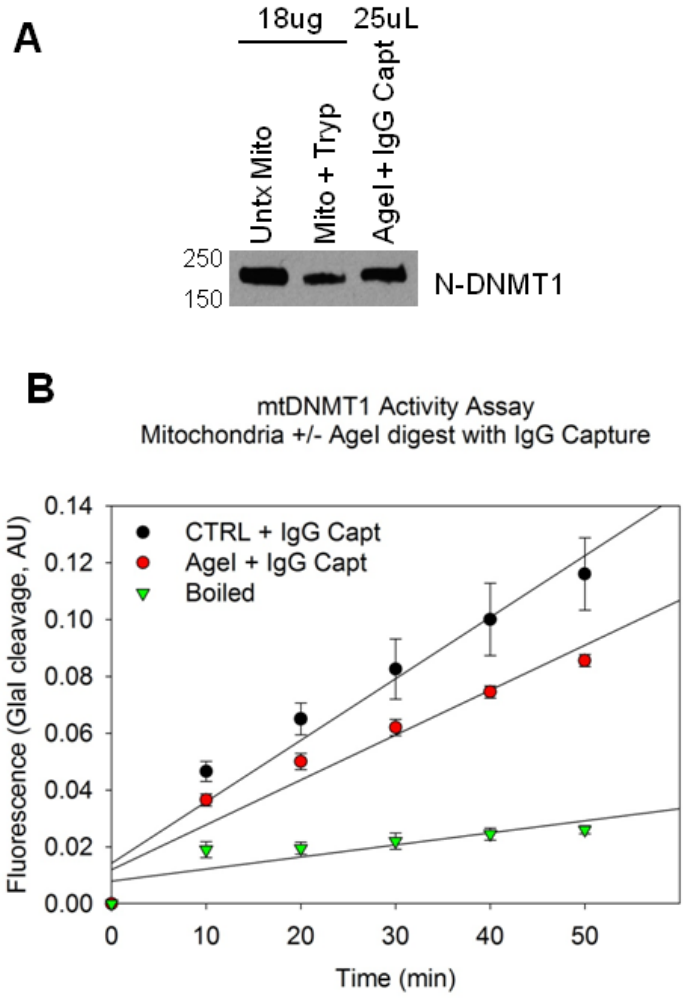


Figure 4-13: AgeI digest of mtDNA allows greater recovery of mtDNMT1-TAP protein, but no increase in enzyme activity. (A) Immunoblot of mitochondrial fractions from HCT116-TAP cells, showing high level of mtDNMT1 in sample subjected to *Age*I digestion and IgG capture of the TAP-tag. The stronger signal in lane 3 demonstrates enrichment for mtDNMT1-TAP in that sample. (B) Despite increased recovery of the mtDNMT1-TAP protein with *Age*I digestion of mtDNA, equal volumes (5µL) of the purified mtDNMT1-TAP enzyme showed no increase in methyltransferase activity.

and that enzymatic cleavage of mtDNA makes the TAP-tag epitope more available for binding by IgG beads, allowing greater recovery of the protein. However, enzyme assays comparing the activities of IgG captured mtDNMT1-TAP that had either been treated with *Age*I or not showed no substantial difference in methyltransferase activity (Figure 4-13B). This indicates that, although digestion of mtDNA allows for a greater recovery of mtDNMT1-TAP, it does not enhance its methyltransferase activity *in vitro*.

5-aza-dC inhibits mtDNMT1 expression and activity *in vivo*

To gain a better understanding of the role DNA methyltransferase activity plays in mitochondria, we wanted to attempt to manipulate the level of DNMT1 protein and thus, DNMT1 activity, in the mitochondria. 5-aza-dC has been shown to be a potent inhibitor of DNA methylation in the nucleus, so we examined its efficacy on mitochondrial DNMT1. HCT116-TAP cells were grown in the presence of the DNMT1 inhibitor, 5-aza-dC, for a total of 48hrs. Cells were harvested and fractionated into whole cell and mitochondrial extracts at both 24hr and 48hr timepoints after treatment, and untreated cells were harvested at 24hrs as a control. Whole cell and mitochondrial lysates were separated by SDS-PAGE on 4-15% gradient gels and probed with antibodies to VDAC as a loading control, and H3K4me3 to demonstrate the absence of contaminating nuclear material in the mitochondrial fractions. As shown in Figure 4-14 (top panel), there is a dramatic reduction in DNMT1 protein in whole cell lysates from 5-aza-dC-treated cells relative to control, and this phenomenon occurs as early as 24hrs post-treatment. Densitometry readings estimate an approximate 80-90% reduction in full-length DNMT1 seen in whole cells treated with the DNMT1 inhibitor. Upon higher exposure, the effect of 5-aza-dC on mtDNMT1 expression can be visualized (Figure 4-14, middle panel). It is evident that, while 5-

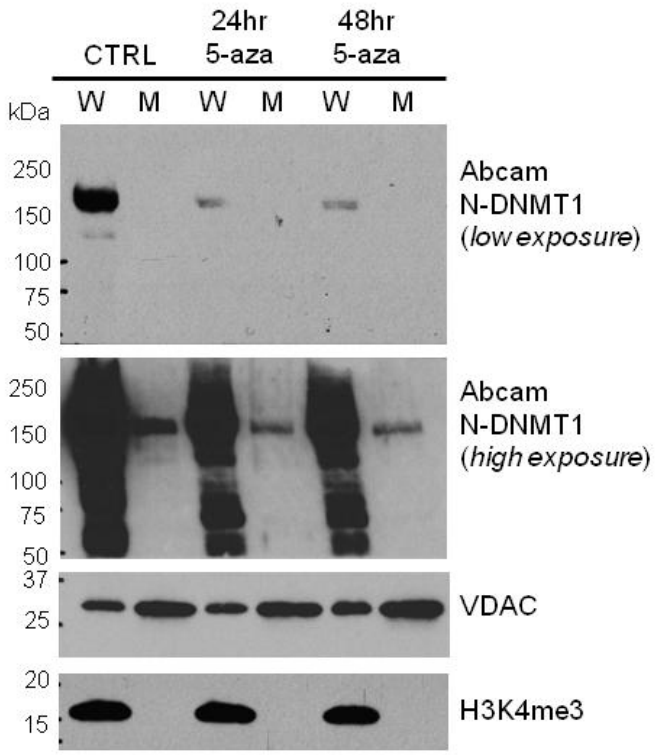


Figure 4-14: Treatment of cultured cells with 5-aza-dC dramatically reduces mtDNMT1 protein. (A) Immunoblots of fractionated HCT116 cells treated with 5-aza-dC for 24 or 48hrs show a striking reduction in DNMT1 expression in whole cell lysates, and a modest ~50% reduction in mtDNMT1 protein. VDAC was used as a loading control; H3K4me3 demonstrates the purity of mitochondrial lysates.

aza-dC induces a clear reduction in mtDNMT1 protein, it is not as potent of an inhibitor in the mitochondrial compartment (~50% reduction in mtDNMT1),

Based on the differential susceptibilities of chromatin-bound and soluble DNMT1 to treatment with 5-aza-dC in the nucleus (141), we postulated that mtDNMT1 might also respond differently to inhibition based on its sub-mitochondrial localization. Therefore we isolated mitochondria from HCT116-TAP cells that had been grown in the presence of 3μM 5-aza-dC for a total of 48 hours. Cells were harvested at 24hr, 48hr and 72hr timepoints after the first exposure to 5-aza-dC to extend our analyses from above. Mitochondria were trypsin-treated, sub-fractionated into S and P fractions and assayed for enzymatic activity under standard conditions. DNA methyltransferase activity from total mitochondrial extracts appeared to be remarkably resistant to the inhibitory effects of 5-aza-dC (Figure 4-15A). However, methylation activity in S and P sub-mitochondrial fractions was modestly affected by 5-aza-dC (Figure 4-15B, C and Table 4-3). At both 24hr and 48hr timepoints, 5-aza-dC-treated S fractions showed little change in methylation activity, relative to the untreated control (Figure 4-15B). By 72hrs, however, a noticeable decrease in DNA methyltransferase activity could be seen in the S fraction of 5-aza-dC-treated cells, compared to control. The insoluble, mtDNA-bound form of mtDNMT1 was affected at an earlier timepoint than the soluble form. Methyltransferase activity in 5-aza-dC-treated P fractions decreased at 24hr and 48hr timepoints, but appeared to rebound slightly by 72hrs (Figure 4-15C, Table 4-3). This dynamic difference in response to 5-aza-dC treatment implies that while mtDNMT1 activity was affected within the first 48hrs, cessation of 5-aza-dC treatment after this timepoint allowed mtDNA-bound DNA methyltransferase to recover and resume more normal methylation activity by 72hrs.

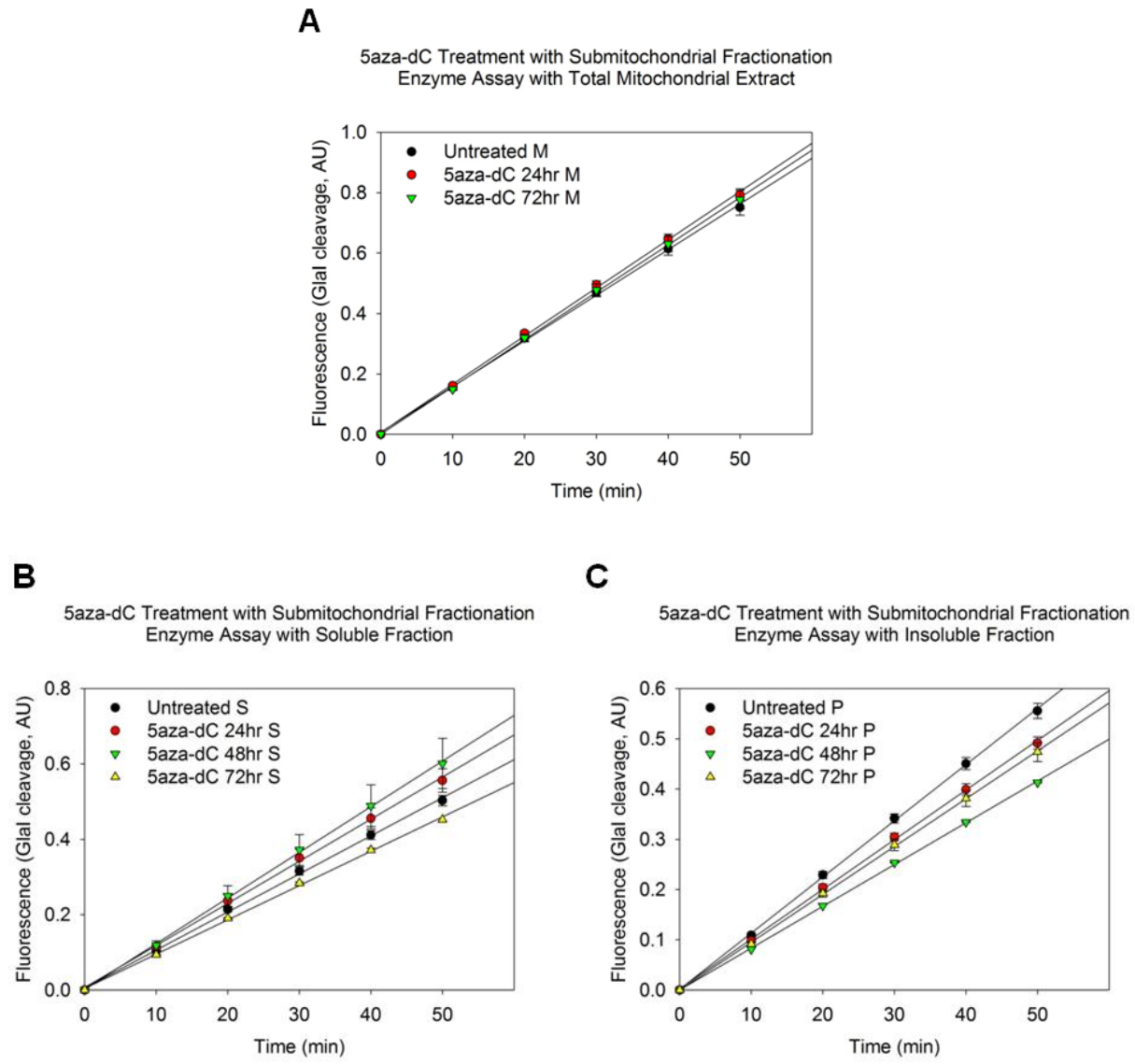


Figure 4-15: Treatment of cultured cells with 5-aza-dC differentially affects mtDNMT1 activity. Total mitochondrial extracts were prepared from untreated control and 5-aza-dC-treated HCT116-TAP cells at 24, 48, and 72hrs after the first exposure to 5-aza-dC. (A) DNA methyltransferase activity in total mitochondrial extracts appears resistant to the inhibitory effects of 5-aza-dC. (B) Soluble mitochondrial extracts show no substantial 5-aza-dC-induced inhibition of activity through 48hrs. By 72hrs, a modest decrease in DNMT activity can be seen as a result of the drug. (C) The insoluble mitochondrial pellet shows more immediate decreases in methyltransferase activity of cells treated with 5-aza-dC for 24 and 48hrs, but by 72hrs the inhibitory effects of 5-aza-dC have worn-off, as methyltransferase activity shows a partial recovery by 72hrs. All assays run using standard conditions (0.8U/well *Gla*I, 240nM Oligo1).

Table 4-3: Rates of DNA methylation in sub-fractionated mitochondria, isolated from 5-aza-dC-treated cells

		5-aza-dC		
	CTRL	24	48	72
Total M	0.0150	0.0159	--	0.0155
S	0.0101	0.0111	0.0120	0.0090
P	0.0111	0.0098	0.0083	0.0095

DISCUSSION

Interaction between mtDNMT1 and mtDNA

Approaching these studies, we hypothesized that mtDNMT1 is functioning in mitochondria much like it does in the nucleus, which if true, would require its interaction with mtDNA. Data from our mitochondrial DNA immunoprecipitation (mtIP) experiments showed that mtDNMT1-TAP indeed binds to various regions of the mitochondrial genome, however this interaction alone was not enough to prove that it was responsible for methylating mtDNA. Without further information, it could have been speculated that perhaps mtDNMT1 was functioning as a scaffold protein, much like TFAM, which is thought to coat the mitochondrial genome, and assist in unwinding of the mtDNA for better accessibility by the transcriptional machinery (10, 12, 14). It is known that DNMT1 can bind a number of DNA molecules that cannot support methylation, including poly(G) and poly(dG)-poly(dC), among others (130). However, the observation that mtDNMT1 differentially binds specific regions within mtDNA in proportion to their CpG content was more indicative of the protein's function in DNA modification. Because cytosines in the context of CpG dinucleotides are the major sites for DNA methylation in mammals, the relative level of mtDNMT1 binding at regions containing these sites was of particular importance. Since the mitochondrial genome does not contain clusters of CpG dinucleotides surrounding gene promoters, termed "CpG islands", we speculate that methylation at even a single CpG could impact mitochondrial transcription of the gene housing that site.

Perhaps even more interesting, our data shows that mtDNMT1 binds to the D-loop control region of mtDNA (Figure 4-8), which is extensively modified by both 5mC and 5hmC residues (Figure 2-3). This region carries the promoters driving transcription initiation of both

heavy and light strands, and represents the assembly site for transcription initiation complexes. The interaction detected between this region and mtDNMT1 further supports a role for mtDNMT1 in directly regulating mitochondrial gene expression, and/or mtDNA replication.

Sub-mitochondrial localization of mtDNMT1

Studies by Patel, *et al.* (141) showing differential partitioning of DNMT1 in the nucleus prompted us to investigate whether mtDNMT1 might also exist in two separate pools, and if so, what relative percentage can be considered "bound" versus "free". Although they gave no estimates as to the proportions of nuclear DNMT1 that are chromatin-bound versus free, Patel, *et al.* (141) stated that the majority of DNMT1 exists within the nucleus in soluble form, which is precisely what we observed in the mitochondria. Utilizing Bradford assay protein quantitations in combination with densitometry to estimate the relative signal for DNMT1 in each fraction on an immunoblot, we can predict that over 8-fold more DNMT1 is soluble and freely available for methylation in the mitochondria, relative to the amount of DNMT1 bound to mtDNA. This may seem counterintuitive for the majority of DNMT1, an enzyme that functions by tightly associating with and modifying DNA, to exist naturally in a state free from DNA. However, a complete understanding of its role in the mitochondria has not yet been realized. It could be that mtDNMT1 not engaged in methylation serves other non-enzymatic functions, such as indirectly influencing the structure of mtDNA or participating in other protein complexes critical to mitochondrial function. In any case, our data suggests that the mitochondrial compartment contains considerably more DNMT1 than it needs for methylation of mtDNA alone.

Why is the methylation activity for the TAP-purified mtDNMT1 enzyme so low in comparison to other samples of reduced purity?

Holding constant the total amount of protein put into each assay, we expected to see the greatest methyltransferase activity from the enzyme preparation with the highest proportion of that total protein represented by mtDNMT1 (the TAP Purif mtDNMT1 sample) (Figure 4-9). It was quite puzzling to observe the opposite effect. At least part of the difference seen in methylation activity between the TAP-purified mtDNMT1 and either the crude or Percoll-purified preparations can be attributed to the fact that only the activity of the mtDNMT1-TAP allele is being measured in the TAP Purif mtDNMT1 sample. If both alleles are expressed equally, then we would expect to see only 50% the level of activity from a single purified DNMT1-TAP allele, compared to two alleles of WT DNMT1. In the crude and Percoll-purified mitochondrial preparations, both copies of mtDNMT1 (WT and mtDNMT1-TAP) are contributing to the assay signal.

However, these results raise a number of other concerns. They could suggest problems with regard to the stability of the purified protein; perhaps the enzyme is unstable in the storage buffer or under the reaction conditions we have chosen. Both the reaction and storage conditions we used are identical to the recommended conditions for the recombinant hDNMT1 enzyme, as described by New England BioLabs. However, these conditions were adapted from cytosolic conditions, and the mitochondrial environment may be quite different with regard to ionic strength and pH, among other factors. Perhaps the mtDNMT1-TAP enzyme would be more functional in reaction conditions that more closely mimic the mitochondria.

Although the entire TAP-tag purification is performed under mild conditions, it is possible that we are purifying the enzyme away from a cofactor or protein binding partner that is

required for more efficient methylation. Since the Crude and Percoll-purified preparations represent total mitochondrial lysate, the necessary factors or accessory proteins would be present to assist with methylation activity in those samples.

It is also possible that restriction endonucleases exist within mitochondria that target a recognition sequence within our fluorescently-labeled oligo substrate, generating an artificially high signal in those reactions. If this is true, then any enzyme preparation made from total mitochondrial lysate (either crude or Percoll-purified) would be affected by active endonuclease activity, which could be misinterpreted as methyltransferase activity. It would also suggest that the signal we detect for either our TAP-purified mtDNMT1 enzyme or the IgG Captured enzyme represents the true methylation activity for mtDNMT1, and the "activity" from total mitochondrial extracts is falsely represented.

Another consideration is that perhaps we are isolating mtDNMT1-TAP that is still so tightly bound to mtDNA it will not release it to bind the artificial Oligo1 substrate. Attempts to digest the mtDNA, either by *Age*I restriction endonuclease or DNase I allowed for capture of more mtDNMT1 protein, as visualized by immunoblot, but no increase in the methylation activity of IgG-captured mtDNMT1-TAP enzyme (Figure 4-12). Likewise, incubation in a high-salt-containing buffer, to strip any bound mtDNA away aided in our efforts to purify mtDNMT1-TAP, but the purified enzyme demonstrated no further increase in methylation activity.

It is possible that mtDNMT1 is not the only methyltransferase being assayed in the crude and Percoll-purified mitochondrial extracts. We showed in Chapter 2 that neither of the *de novo* methyltransferases, DNMT3a and DNMT3b, could be detected in highly purified mitochondrial lysates, and we know from bioinformatics analyses that neither of these enzymes contains an obvious mitochondrial targeting sequence. However, we have not rigorously excluded their

possible presence in mitochondria. It is possible that either or both of these enzymes has a mitochondrial counterpart, and could be contributing to the assay signal we see from whole mitochondrial extracts. One characteristic of the enzyme assay used to measure mitochondrial methylation activity is that it employs a hemi-methylated oligo substrate, which is highly preferred (3-50-fold) over un-methylated DNA substrates by the maintenance enzyme, DNMT1 (131, 132). Therefore, we would not expect either DNMT3a or DNMT3b to exhibit a substantial level of methyltransferase activity on an unfavorable template. However, if both *de novo* and maintenance DNMT enzymes are involved in mtDNA methylation, then it is possible that they share redundant functions and coordinate methylation of DNA substrates as they do in the nucleus.

Pharmacologic inhibition of mtDNMT1 activity

Early studies with the nucleoside analog 5-aza-2'-deoxycytidine (5-aza-dC) demonstrated that treatment of cells with this compound led to a loss of extractable DNA methyltransferase activity (142). Data described throughout this Chapter consistently demonstrated the presence of methyltransferase activity within mitochondria, although the exact details of the mechanism of mtDNA methylation have not been elucidated. We hypothesized that by examining the 5-aza-dC-mediated inhibition of mtDNMT1 we may gain a better perspective of how the enzyme functioned under non-inhibitory conditions.

The success of this experiment was based on a couple of assumptions: first, that 5-aza-dC can gain access to the mitochondria, and second, that 5-aza-dC can be phosphorylated and become incorporated into mtDNA. It was not unreasonable to assume either, as mitochondria are capable of replication and repair of their own DNA. As our data shows, 5-aza-dC effectively

reduces the amount of mtDNMT1 protein seen by immunoblot, although not nearly to the same extent as its effect on total DNMT1 in whole cells (Figure 4-14).

There are thought to be two mechanisms of action for 5-aza-dC in the nucleus: 1) covalent trapping of the DNA-bound form of DNMT1, and 2) proteasomal degradation of the free form of DNMT1 (141). A recent study described differential effects of 5-aza-dC based on whether the enzyme was soluble or insoluble in the nucleus (141). The majority of DNMT1 was found to be soluble and more susceptible to degradation than the chromatin-bound form.

A sub-fractionation of mitochondria demonstrated that active mtDNMT1 is found in two separate pools within the mitochondrial compartment: the soluble fraction, and the insoluble mtDNA-bound fraction (Figure 4-7). We extended these findings to our 5-aza-dC studies, to see if the state of mtDNMT1 affected its susceptibility to inhibition by 5-aza-dC. Surprisingly, although there was a clear effect on mtDNMT1 protein following exposure of cells to 5-aza-dC, the effect on methyltransferase activity could not be observed by looking at total mitochondrial extracts (Figure 4-15A). Inhibition of DNMT activity by 5-aza-dC was only apparent after separating the mtDNA-bound and soluble forms of mtDNMT1 (Figure 4-15 B, C and Table 4-3). Interestingly, the soluble (S) and insoluble (P) forms of mtDNMT1 responded somewhat differently to treatment with 5-aza-dC, which may suggest two possible mechanisms of action for 5-aza-dC in the mitochondria as well as the nucleus. The S fraction exhibited an effect at a later timepoint than the P fraction, as a reduction in activity was not observed until 72hrs after the first exposure. In contrast, the effect of 5-aza-dC on mtDNA-bound mtDNMT1 was more immediate, with noticeable decreases at 24 and 48hrs post-treatment, whereas by 72hrs after the first exposure, methylation activity seemed to be rebounding to more normal levels. Overall, it was evident that, although 5-aza-dC did appear to induce differential effects on mtDNMT1

activity dependent on whether the enzyme was bound to mtDNA or freely soluble, the drug was not nearly as efficient in mitochondria as it was in the nucleus. This could be representative of the compound's low-level incorporation into mtDNA, or the overall low levels of methylation activity within mitochondria. In any case, we cannot conclude that 5-aza-dC would be a useful tool for inducing changes in either the mtDNA methylation profile or the mitochondrial transcription program.

Chapter 5: Perspectives

I. Summary

Cytosine methylation of mitochondrial DNA (mtDNA) was first described several decades ago, but neither the mechanism generating this modification nor its functional significance was known. Using MeDIP analyses, we show that mtDNA contains not only 5-methylcytosine (5mC), but also 5-hydroxymethylcytosine (5hmC), suggesting that previous reports likely underestimated the degree of epigenetic modification within the mitochondrial genome. As the mitochondrial genome does not encode its own methyltransferase, we questioned whether a nuclear-encoded enzyme might be responsible for initiating these modifications. We found that an isoform of the most abundant mammalian methyltransferase, DNA methyltransferase 1 (DNMT1) translocates to mitochondria, driven by an in-frame mitochondrial targeting sequence (MTS) located upstream of the nuclear DNMT1 translational start site. This MTS is highly conserved across mammalian species, and directs GFP to the mitochondria. Using a gene targeting approach, we inserted a tandem affinity purification (TAP) tag into the endogenous locus of a single human DNMT1 allele by homologus recombination. We have used this cell line to show that mtDNMT1 specifically binds mtDNA in a manner that is proportional to CpG density, and that mtDNMT1 exhibits CpG-specific methyltransferase activity *in vitro.* Mitochondrial DNMT1 is resistant to trypsin-treatment of intact mitochondria, but susceptible to pharmacologic inhibition by the nucleoside analog 5-aza-2'-deoxycytidine (5-aza-dC). Mitochondrial DNMT1 is also responsive to a number of factors known to be intimately

involved in mitochondrial function. NRF1 and PGC1α, transcription factors that activate nuclear-encoded mitochondrial proteins in response to oxidative stress, up-regulate expression of mtDNMT1. Loss of p53, a tumor suppressor gene known to help control mitochondrial metabolism, also results in a striking increase in mtDNMT1 expression, and this up-regulation of mtDNMT1 appears to modify mitochondrial transcription in a gene-specific fashion. Our data suggests roles for mtDNMT1 in both the establishment and maintenance of cytosine methylation, which we propose is involved in the regulation of mitochondrial transcription.

II. Potential utility of the DNMT1-TAP cell line

We generated a unique cell line that carries a TAP tag fused to the C-terminus of a single endogenous human DNMT1 allele. We opted to target the genomic DNMT1 locus, rather than ectopically express a tagged cDNA construct, to avoid issues of overexpression from an unnatural promoter, and artifactual interactions or off-target effects as a consequence of such overexpression. The DNMT1-TAP targeting construct was designed against the C-terminus to allow for expression of the TAP-tagged allele under control of its native promoter and in an endogenous context. Among our reasons for creating the HCT116-TAP cell line using this approach was to have an endogenously-expressed DNMT1-TAP allele that could easily be tracked to its subcellular compartments using an antibody specific for the TAP-tag, instead of the commercially-available antibodies for DNMT1, which are somewhat promiscuous and consistently generate complicated patterns of multiple protein bands. We also desired a system whereby we could isolate DNMT1-TAP from its subcellular compartments and ask questions regarding its methyltransferase activity and its protein-protein interactions without any prior knowledge as to their identity. The strategy we chose to generate the HCT116-TAP cell line has

been used successfully by several groups, to tag and isolate a variety of proteins in both yeast and mammalian cells (107, 108, 113, 114), and has proven an enormously useful tool in understanding the functions and protein partners of those proteins. Thus, we were confident that a DNMT1-TAP allele would be a valuable reagent for our studies on mitochondrial DNMT1, and has already been used in a number of applications to determine mtDNA binding and catalytic activity of mtDNMT1.

However, a potential complication in any experiments that analyze the function or interactions of nuclear DNMT1 is our observation that the DNMT1-TAP allele is expressed at levels 6-8-fold higher than the WT allele at the mRNA level, and DNMT1-TAP appears to be up-regulated at the protein level with two commonly-used DNMT1 antibodies. If this over-expression is not simply an artifact of the antibodies used, then the over-expression of DNMT1 could severely impact the utility of this cell line in determining protein-protein interactions both qualitatively, due to a risk of false-positive or artificial interactions, and quantitatively by disrupting normal stoichiometric ratios of protein complex components. The HCT116-TAP cell line was derived from human colon carcinoma cells, so it exists as a tumor cell line already, which has been shown to exhibit over-expression of DNMT1 and contain regions of hypermethylation (81, 82). In addition to the characteristics of the parental cells, the DNMT1-TAP allele may be expressed as much as 6-8-fold over the WT allele, thus likely exacerbating the already aberrant levels of DNMT1 and methylation within this cell line.

However, the mitochondrial form of DNMT1-TAP does not appear to be overexpressed relative to WT DNMT1 at either steady-state protein or mRNA levels (Figures 3-16 and 3-17), and mitochondria from HCT116-TAP cells demonstrate rates of methyltransferase activity comparable to non-tagged HCT116 cells (Figure 4-11). Thus, while the HCT116-TAP cell line

may only cautiously be used to examine factors influencing DNMT1 in whole cells or nuclei, it should be an invaluable reagent for understanding the function, protein-protein interactions, and post-translational modifications of DNMT1 in the mitochondrial compartment, as discussed below.

Among the distinct advantages of the HCT116-TAP cell line is that it allows for purification of proteins in native complex with DNMT1-TAP. This can be done from whole cells, as shown in Chapter 3 (Figures 3-12 and 3-18), or the analysis can be reduced to focus on protein-protein interactions within a particular sub-cellular compartment, such as the nucleus or the mitochondria. We attempted to purify mtDNMT1-TAP from isolates of highly purified mitochondria and recovered a sufficient amount of mtDNMT1-TAP to analyze by enzymatic assay, but we observed a poor recovery of the full-length protein after TAP-tag purification or IgG Capture, as only a faint band could be visualized by immunoblot with a DNMT1 antibody or silver stain (Figure 5-1). We determined that we can use restriction enzyme digestion and/or DNaseI treatment of mtDNA to capture more full-length mtDNMT1-TAP in an immunoprecipitation with IgG beads (Figure 4-12A). Thus, these methods can be used to pull-down mtDNMT1-TAP and begin to ask what other proteins co-purify with it. TAP-tag purification, in combination with non-biased mass spectrometry (MS) or direct immunoblotting experiments with antibodies to proteins presumed to interact, can be a powerful tool to identify interacting protein partners of mtDNMT1-TAP.

Likely candidates for interacting partners include TFAM, POLG and mTERF1, which could all be involved in complexes with mtDNMT1 at sites of mtDNA replication and active transcription. Antibodies against all three proteins have been validated for application in both immunoblot and immunoprecipitation experiments. The mTERF1 antibody, in particular, could

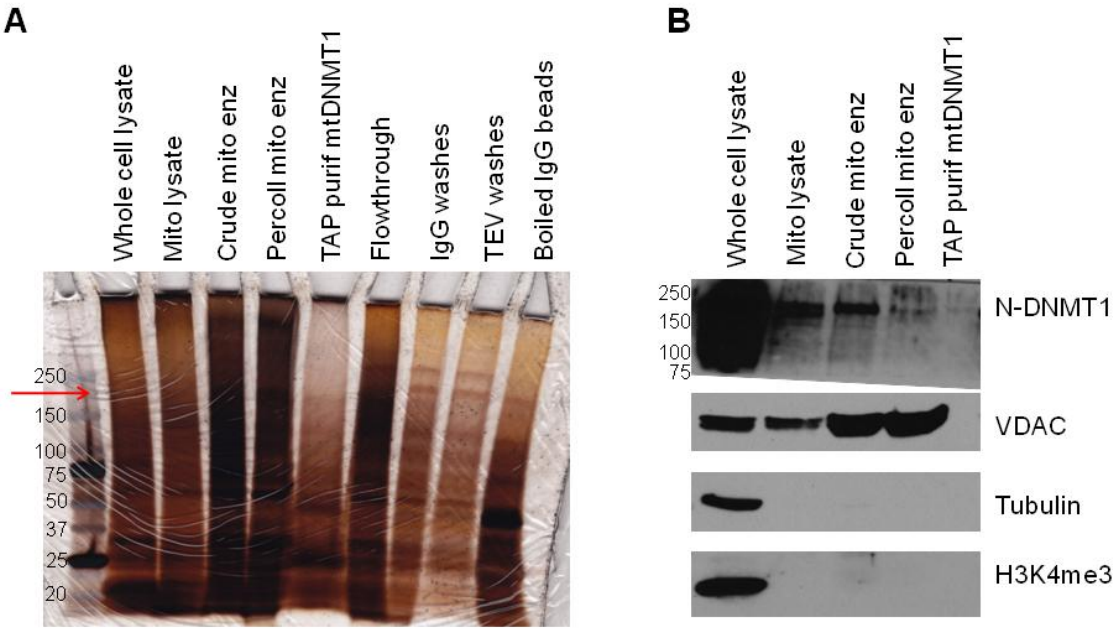


Figure 5-1: TAP-tag purification of mtDNMT1-TAP. (A) Purification fractions were resolved on a 4-15% gradient SDS-PAGE gel, and protein bands were visualized by silver stain. Only a faint band is visible for full-length mtDNMT1 at ~185kDa (red arrow) in lane 5 for the purified elution fraction, indicating poor yield of the purified mtDNMT1-TAP. (B) Immunoblots of purification fractions probed with an N-DNMT1 antibody only faintly detect full-length mtDNMT1 in the purified elution fraction (lane 5). Antibodies to compartment-specific markers were used to demonstrate the purity of subcellular fractions: VDAC, mitochondrial marker; tubulin, cytosolic marker; H3K4me3, nuclear marker.

be used to start teasing apart the gene-specific effects observed in mitochondrial transcription upon over-expression of mtDNMT1. A co-immunoprecipitated sample of mtDNMT1-TAP pulled-down with IgG beads or the TAP-tag antibody could be resolved by SDS-PAGE and probed with the mTERF1 antibody to observe if an interaction is first present in a situation where mtDNMT1 is not overexpressed. Then, experiments to perturb the mTERF1-mtDNMT1 interaction with analyses of downstream mitochondrial transcriptional effects can be performed as described in more detail below, in a proposed model for transcriptional control in mitochondria. Similarly, co-immunoprecipitation experiments with the TFAM antibody could further implicate mtDNMT1 as being involved in transcription, and extend the correlative observations we made in Chapter 2. However, an association with TFAM could also suggest that mtDNMT1 is involved in overall mtDNA architecture, influencing the structural unwinding of mtDNA during transcription, as a role secondary to its methylation activity. If an interaction with POLG were detected, it could suggest a replicative function for mtDNMT1 through participation with a mtDNA replication complex, or simply, that it associates with sites of active mtDNA synthesis as a consequence of its methyltransferase activity, much like it has been shown to localize to replication foci in nuclear chromatin.

As an extension of the studies proposed above, the DNMT1-TAP allele purified from both whole cells and within mitochondria was shown in Chapters 3 and 4 to exhibit catalytic activity *in vitro*. This is a critical observation, because it implies that this cell line can be used to examine factors that may influence methyltransferase activity, in either the nuclear or mitochondrial compartments. Although the activity of DNMT1 *in vivo* is thought to require assistance from accessory proteins, such as UHRF1, to localize properly to hemi-methylated CpG sites within DNA for methylation fidelity, *in vitro* enzymatic reactions seem to proceed

normally in the absence of these factors. Thus, using an *in vitro* system to draw conclusions about how methylation occurs in an *in vivo* context might lead one to conclude that DNMT1 is capable of methylating DNA on its own. However, among some of the most puzzling data discussed throughout this dissertation was the observation that a highly purified preparation of mtDNMT1-TAP enzyme exhibits lower activity than total mitochondrial extracts against the same hemi-methylated DNA substrate. Attempts to make the mtDNMT1-TAP allele more available and to strip away any bound mtDNA for higher recovery of activity proved futile. Thus, it is possible that a protein binding partner or other co-factor, removed during purification, might be necessary for more efficient methylation by mtDNMT1. In this instance, the *in vitro* enzymatic assay, in combination with a means of systematically re-constituting the system, could be essential tools in identifying what is missing from our original mtDNMT1 catalytic assays. Perhaps there is a protein in the mitochondria that plays an analogous role to UHRF1 in the nucleus, and its association with mtDNMT1 is required for methylation activity, even in an *in vitro* setting. If that is the case, then we would expect to see a spike in methyltransferase activity when that factor is incrementally added back to the reaction. An alternative possible explanation for loss of activity following purification of the TAP-tagged mtDNMT1 is that the TAP-tag interferes with catalytic activity. The TAP-tag insertion adds considerable sequence at the C-terminus of DNMT1, which could also introduce unnatural secondary or tertiary structure, thereby limiting accessibility of the active site of the enzyme to its target cytosine substrate. However, a direct comparison of methyltransferase activity in mitochondrial extracts isolated from both non-tagged and TAP-tagged HCT116 cells did not show a remarkable difference. If the TAP-tag was interfering with activity, we would expect to see an approximate 50% reduction

in the level of activity measured in HCT116-TAP cells, with only one WT allele, compared to non-tagged cells, with both WT alleles.

Purification of mtDNMT1-TAP from mitochondria, and further characterization of its co-purified proteins could also shed light on the identities of the other smaller proteins we detect with DNMT1 antibodies in the mitochondria. A 60kDa band is seen with an N-terminal DNMT1 antibody and a 30kDa immunoreactive species is visualized with the C-terminal antibody (Figure 3-16). We are still unsure whether these proteins are proteolytically processed products of the full-length mtDNMT1 or non-specific proteins, detected due to promiscuity of the antibodies. Although mass spectrometry usually cannot be used to obtain the full sequence of a protein, it is more valuable instead as an identification tool; MS could provide enough information within discrete peptide sequences to predict both the identities and the precursor proteins for those smaller species. Alternatively, N-terminal sequencing could be performed to obtain full sequences for the unknown protein bands co-purified with mtDNMT1. Knowing the identities and compositions of these smaller protein species would help to explain, for instance, why we see no change in the level of the 60kDa mitochondrial protein species upon loss of p53, while the full-length mtDNMT1 is dramatically up-regulated (Figure 2-12). If this protein species is indeed a proteolytic product of the full-length enzyme but still contains the N-terminal epitope, one would expect it to be equally increased in cells lacking p53, unless the processing peptidase within mitochondria (MPP) is saturated, thereby limiting digestion of the full-length enzyme.

In line with the studies described above to reveal the identities of the smaller protein species in mitochondria, the HCT116-TAP line could be used to assay their methyltransferase activities. If those smaller species represent proteolytically-processed forms of the full-length mtDNMT1, then we can already speculate what roles they may play and whether or not they

might contain catalytic activity based on what is known about the functional domains of DNMT1 in the nucleus. The epitope-specific antibodies against DNMT1 used for detection by immunoblot are informative, such that they can be used to predict which conserved motifs might be present in a product cleaved from the N- or C-terminus. For instance, the C-terminal catalytic domain that is so highly conserved across both eukaryotic and prokaryotic CpG methyltransferases (Figure 1-5) contains amino acid sequence that, itself, would produce a protein approximately 60kDa in size. This means, however, that the 30kDa band detected with the C-terminal antibody is not likely to demonstrate catalytic activity. Similarly, although the 60kDa band seen on immunoblots with an N-terminal DNMT1 antibody is of sufficient size to contain the catalytic domain, its detection with an N-terminal epitope suggests that it likely only comprises a region within the regulatory domain, and thus, would also not be expected to exhibit methyltransferase activity. These hypotheses could be tested using the HCT116-TAP cell line by first fractionating mitochondrial extracts by size-exclusion chromatography, and then performing additional TAP-tag purification on those fractions. Because the TAP-tag is inserted into the C-terminus, we would only be able to purify the 30kDa band detected with a C-terminal antibody, but the additional size exclusion step would allow us to distinguish between methyltransferase activity coming from the smaller protein species versus the full-length enzyme. In the nucleus, it has been shown that the catalytic domain of DNMT1 is not functional on its own; rather, a generous portion of the N-terminal regulatory domain is required for efficient methyltransferase activity (74). If this observation holds true in the mitochondria, then it is likely that only the full-length enzyme (after cleavage of the MTS) is responsible for methylation of mtDNA.

As discussed in the introduction, nuclear DNMT1 is known to be regulated by post-translational modification (PTM) at several key residues, but this area remains largely

uninvestigated. Nothing is known about post-translational modification of mtDNMT1, so the potential for these mechanisms to control the function of mtDNMT1 begs exploration. Again, because the HCT116-TAP cell line carries a TAP-tag that has been inserted into the endogenous DNMT1 locus, the resultant DNMT1-TAP fusion protein is expected to be subject to regulation by PTMs just like the WT allele. It is well-established that mitochondria are highly responsive to conditions of oxidative stress, which generates cysteinyl-NO (SNO) adducts on many proteins (146), and triggers phosphorylation cascades (147), thus impacting OXPHOS pathways and mitochondrial biology in general. If PTMs were identified for mtDNMT1, the HCT116-TAP cell line could be the ideal reagent to measure changes in the profiles of PTMs to mtDNMT1 as a function of oxidative stress. This would broaden our findings that mtDNMT1 is responsive to stimulation by NRF1 and PGC1α (Chapter 2), which themselves are up-regulated by oxidative stress, and provide direct insight into how mtDNMT1 itself is controlled and thereby elicits downstream effects on mitochondrial physiology.

Just as the studies above propose investigating the alterations in specific PTMs of mtDNMT1 under conditions of oxidative stress, the HCT116-TAP cell line can be used to study changes in protein-protein interactions, methylation activity and presumed proteolytic products of mtDNMT1 under various conditions. It is relatively easy to induce oxidative stress in cultured cells; they can be grown in a hypoxic chamber under conditions of 2% O_2, in tissue culture flasks with non-vented caps, in the presence of cobalt chloride or through chronic exposure to alcohol. Several of these methods are currently being employed by a graduate student, Joyce Balinang, in our laboratory, and preliminary results show dynamic alterations in levels of mtDNMT1 under hypoxic conditions, recapitulating the effect shown in Figure 2-11. Induction of oxidative stress within the HCT116-TAP cell line, isolation of mitochondria and subsequent purification of the

mtDNMT1-TAP allele would provide the cellular reagents necessary to assess changes in mtDNMT1 levels, activity and co-purified protein partners in a single experiment.

We also used the HCT116-TAP cell line in mitochondrial DNA IP experiments to demonstrate that mtDNMT1-TAP binds to mtDNA in a CpG-dependent manner (Figure 4-8). These assays should also be performed in different cellular contexts, such as oxidative stress, to determine whether changes in mtDNMT1 protein levels seen by immunoblot correlate with an altered amount of mtDNMT1 bound to mtDNA. Although we would expect increases in mtDNMT1 protein levels to coincide with increased binding of mtDNMT1 to sites along mtDNA, and therefore increased mtDNA modification, this highlights an important point: it may prove challenging to distinguish between effects caused by changes in mtDNMT1 protein levels and those caused by changes in mtDNMT1 function. The mitochondrial form of DNMT1, as in the nucleus, performs multiple tasks: it binds mtDNA, and exhibits methyltransferase activity. However, it may serve other roles unique to its compartmentalization within mitochondria. We will therefore need to rely on other methods, such as MeDIP, hydroxy-MeDIP, enzyme assays and transcriptional analyses to precisely tease-apart the many possible roles mtDNMT1 could be playing to impact mitochondrial function. In addition to methylation-mediated effects on mitochondrial transcription, other potential functions include serving as a scaffold protein, similar to the role of TFAM (10, 14), binding mtDNA, and recruiting or blocking the interaction of specific mitochondrial transcription factors or the transcriptional machinery with mtDNA.

III. Proposed model for transcriptional control in mitochondria

The mitochondrial transcriptional studies described in Chapter 2 led us to question whether the changes in mitochondrial transcription patterns observed were strand-specific or

gene-specific, and whether up-regulation of mtDNMT1 or loss of p53 was responsible for those effects (Figure 2-13 and 2-14). It is clear that this issue needs to be addressed in a more defined cellular context, where only mtDNMT1 activity has been altered and the confounding effects of p53 status can be avoided.

Because the 5' region that distinguishes the mitochondrial and nuclear forms of DNMT1 is so small (303bp in human, 159bp in mouse), and the precise transcription start site of mtDNMT1 remains to be located, RNAi strategies to knock down expression of only the mtDNMT1 isoform are not feasible. Thus, we have designed approaches to both knock-out and overexpress mitochondrial DNMT1. The knockout strategy involves targeting of both alleles to modify exon 1 by adding two termination codons between ATG2 and ATG3 (Figure 2-6). The targeting construct contains either hygromycin or zeocin resistance cassettes, allowing conditional deletion of each allele of mtDNMT1 individually. In similar conditional knockouts of total DNMT1 in HCT116 cells, En Li and colleagues (148) reported that total loss of DNMT1 function led to mitotic catastrophe, characterized by G2 arrest, severe mitotic defects and cell death. If this striking cellular reponse has a mitochondrial component, then it would suggest an essential role for mtDNMT1 in cell survival. However, if the targeted cells were able to escape these effects, then they could be useful in determining what effects loss of mtDNMT1, and therefore loss of mtDNA methylation, has on mitochondrial transcription and function.

In an alternative approach, we have generated all of the reagents necessary to induce overexpression of only the mitochondrial form of DNMT1. These include CMV-driven expression constructs containing DNMT1 sequences starting at: ATG1, which would express both mitochondrial and nuclear forms; ATG3, which expresses exclusively the nuclear form; and a ΔATG3opt construct, in which ATG3 has been mutated to alanine, and the Kozak sequence

surrounding ATG1 has been optimized, permitting high overexpression of only the mitochondrial isoform. These expression constructs have been transferred into lentiviral vectors for efficient transduction of even the hardest-to-transfect cell lines.

Either of these genetic approaches should provide an adequate system for more specifically altering mtDNMT1 levels, enabling us then to look at direct consequences, including mtDNA methylation and hydroxymethylation profiles, and changes in mitochondrial transcription. Based on our initial findings, we predict that altering levels of mtDNMT1 will have effects on mitochondrial transcription, particularly visible on the heavy strand, where there are several protein-coding ORFs and two different promoters driving transcription. We hypothesized from our original experiments that the increase in ND1 expression on the heavy strand could be a result of interference with proper termination from HSP1, introduced either by methylation of the terminator binding site, or through direct protein-protein interaction of mtDNMT1 bound at that site (Figure 2-16). Using cell lines either lacking or overexpressing only mtDNMT1, we can perform mtIP experiments to look for changes in the residence of mtDNMT1 and/or mTERF1 at the termination site, and whether transcription of ND1 and the rRNA genes is subsequently altered.

We have obtained several other cell lines that will be instrumental in our analysis of functional outcomes of mtDNMT1, methylation and hydroxymethylation in mitochondria. HCT116 cell lines carrying hypomorphic alleles of DNMT1 (DNMT1-/-), null alleles for DNMT3b (DNMT3b-/-), or double knockout (DKO) were received from Dr. Bert Vogelstein (Johns Hopkins University). The hypomorphic DNMT1 alleles were generated by replacing exons 3-5 with a hygromycin resistance cassette, however, alternative splicing from exons 1-6 resulted in low-level production of a truncated protein that retains some hemimethylation activity

(66, 67). Although enough DNMT1 is still present in the DNMT1-/- hypomorphs to maintain approximately normal levels of methylation, the value of these cell lines will be in determining whether DNMT1 is solely responsible for the methylation observed in mitochondria. We have not been able to justify why we observe greater methyltransferase activity in crude mitochondrial preparations, in comparison to a more pure mtDNMT1-TAP enzyme (Figure 4-9). One possible explanation for this is that mtDNMT1 may not be the only methyltransferase that gains access to the mitochondrial compartment. Although we could not detect either of the *de novo* methyltransferases, DNMT3a or DNMT3b, in the mitochondrial fractions of cells probed by immunoblot, and we cannot locate obvious mitochondrial targeting sequences within the genomic sequence for either enzyme through bioinformatics, it will be important to know whether they are present or absent in mitochondria with certainty. If one or both of these enzymes is present in mitochondria, then genetic deletion of either DNMT3a or DNMT3b, or both together (3a-/- 3b-/- DKO), would induce changes in 5mC modification of mtDNA as assessed by MeDIP, and mitochondrial methyltransferase activity as measured by enzymatic assay. Furthermore, if one of the *de novo* methyltransferases is found to translocate to mitochondria, then distinct roles for *de novo* and maintenance methylation, and preference for hemi-methylated versus unmethylated DNA substrates, among other properties, will need to be examined in mitochondria similar to studies that have been performed for the nuclear DNMTs.

Mouse embryonic stem (ES) cells containing deletions in DNMT1 exons 30-32, which encode the PC and ENV motifs required for catalytic activity, were obtained from Drs. Masaki Okano and En Li. (58). The J1 wild type ES cells were used to generate heterozygous (C/+) as well as homozygous deletion (C/C) ES cell lines. Prashant Thakkar in our laboratory has been working closely with these three ES cell lines, and has preliminary data showing an approximate

50% reduction in the level of mtDNMT1 in C/+ cells compared to WT J1, and the C/C cells show no detectable levels of mtDNMT1. These cell lines will be critical in a more stringent analysis of mtDNMT1 function, as it relates to 5mC and therefore 5hmC modification in mitochondria. Data generated through MeDIP and hydroxy-MeDIP analyses of these cell lines should indicate whether the 5hmC modification in mtDNA depends on the presence of 5mC, or if 5hmC can arise independently through some other mechanism, as discussed in Chapter 2. This distinction is critical to the way we think about how the processes of methylation and hydroxymethylation are themselves regulated, and how they can be used to modify mitochondrial transcription.

If 5hmC is derived from 5mC, then the enzyme responsible for its conversion would be extremely important to overall levels of mitochondrial epigenetic modification, potentially through demethylation of mtDNA, and would probably play roles in influencing mtDNA structure as well as mitochondrial transcription. Prashant Thakkar in our laboratory is already examining whether TET1, the most abundant of the Tet proteins in ES cells and responsible for conversion of 5mC to 5hmC in the nucleus, might also have a mitochondrial component. If a mitochondrial TET1 protein is identified, it will be essential to understand whether it interacts with mtDNMT1, and how the function of the TET1 enzyme is balanced by mtDNMT1 methylation to coordinate effects on mitochondrial gene expression. As described in Chapter 1, the patterns for 5mC and 5hmC modifications have been mapped for the nuclear genome, and it appears as though they exhibit opposite distributions with regard to CpG-rich (low 5mC, high 5hmC) and CpG sparse (high 5mC, low 5hmC) regions. Thus, it almost seems as though the high presence of 5hmC within CpG-rich promoters may confer some degree of protection against the 5mC modification, and subsequent methylation-mediated silencing of the associated genes. If

this “competition” between 5mC and 5hmC for the same CpG sites also occurs in mitochondria, then factors dictating which modifications are present within mtDNA could serve as critical regulators of mitochondrial gene expression.

Because CpG dinucleotides are under-represented and generally sparse throughout mtDNA, it is possible that methylation of even a single residue could impact mitochondrial transcription or binding by proteins associated with mtDNA. It is therefore particularly important that we more precisely map the distribution of 5mC modification in the mitochondrial genome. MeDIP analyses paired with qPCR can report relative quantitations for the level of 5mC modification within a specific amplicon, but it cannot provide information as to the methylation status of a particular site. There is a distinction to be made between overall levels of mtDNA methylation and methylation status at a single cytosine residue; we will need to perform next-generation sequencing of bisulfite-modified mtDNA to determine the methylation status of mtDNA at a single nucleotide resolution. We have already optimized the conditions necessary to achieve complete bisulfite conversion, and preliminary reads of the mitochondrial genome detect 5mC modifications within D-loop regulatory region of mtDNA, which houses the promoters for bi-directional transcription as well as the origin of replication for heavy strand synthesis (Figure 1-2). This data complements what we observed by mtIP showing mtDNMT1 bound in the D-loop region, and by MeDIP analysis with the 5mC antibody (Figure 2-3). Additionally, if 5mC is a precursor for 5hmC, it could also further support data from the hydroxy-MeDIP and β-gt assays, showing 5hmC modification in the D-loop (Figure 2-3 and Figure 2-5).

IV. mtDNMT1 as a possible therapeutic target

Mitochondrial dysfunction has been implicated in a wide variety of human diseases, including heart disease, neurological disease, diabetes, aging and cancer. Central to each of these disorders is an alteration in the production of cellular energy, mediated by mitochondrial glycolysis and oxidative phosphorylation. Mitochondrial energy production could not be achieved without assistance from the nuclear genome, which encodes the majority of proteins required for OXPHOS complex formation, mitochondrial biogenesis, and mitochondrial DNA transcription, which produces essential components of the respiratory chain complexes.

Two of the major nuclear-encoded factors that exert control over mitochondrial function are NRF1 and PGC1α. These transcriptional co-activators stimulate transcription of nuclear-encoded genes with protein products destined for import into mitochondria as part of a cellular response to changing environmental conditions. Our data demonstrates that among the nuclear-encoded mitochondrial proteins regulated by NRF1 and PGC1α is the novel mitochondrial isoform of the maintenance mammalian DNA methyltransferase, DNMT1. Roles for DNMT1 in methylation of nuclear DNA and methylation-mediated regulation of nuclear gene expression have been well-established. Our data showed that up-regulation of mtDNMT1 was associated with distinct changes in the mitochondrial transcriptional program, which suggests that mtDNMT1 may serve to regulate mitochondrial gene expression through a mechanism analogous to its role in the nucleus.

Pharmacologic inhibition of DNA methyltransferase activity through incorporation of the nucleoside analog 5-aza-2'-deoxycytidine (5-aza-dC) into DNA was first demonstrated in cultured cells (39, 40, 142), and has since been approved for clinical use in treating myelodysplastic syndrome and acute leukemias. Combination clinical trials are currently

underway, using 5-aza-dC paired with histone deacetylase inhibitors and interleukin-2 in the treatment of renal cell carcinoma as well as melanoma (149). In the nucleus, DNMT1 protein and activity are effectively shut down by treatment with 5-aza-dC, and genomic methylation levels have been reported to decrease by 80% (142). We examined the effects of this compound on mtDNMT1, and observed only a modest inhibition of DNMT1 in the mitochondrial compartment, relative to its substantial effect on DNMT1 in whole cells (Figure 4-14). Thus 5-aza-dC appeared to be relatively ineffective at inhibiting mtDNMT1 activity (Figure 4-15). Studies using other nucleoside analogs of cytidine, including 2'-deoxy-2',2'-difluorocytidine (dFdC, gemcitabine), report that the phosphorylated form (dFdCTP) is incorporated into mtDNA with a 432-fold lower efficiency than dCTP (150). Additionally, these bases were interpreted as mtDNA lesions, causing POLG to pause and incorporate downstream nucleotides much more slowly, and with 2-fold less fidelity at pause sites (150). Although no previous reports have used 5-aza-dC to inhibit mtDNA methylation, our results likely suggest that this compound, like dFdC, is also not incorporated efficiently into mtDNA. Further experimentation is needed to confirm this theory.

Although mtDNMT1 does not appear to be as sensitive to the effects of the mechanism-based inhibitor, 5-aza-dC, mtDNMT1 may still serve as a relevant therapeutic target using compounds against one of its other proposed functions, such as mtDNA binding. Alternatively, 5-aza-dC could be used in combination therapy perhaps with analogs of 5hmC, which could be effective if mtDNMT1 turns out to be directly involved in generating the 5hmC modification. The possibility of targeting mtDNMT1 as a means of regulating mitochondrial function is certainly present, but it will be important for future studies to further discern the specific roles of mtDNMT1 in the mitochondria so that their therapeutic potential can be realized.

V. Conclusions

Mitochondria are intimately involved in normal cell function as they are responsible for the production of the majority of cellular energy required for processes including cell signaling, differentiation, and control of cell cycle and cell growth (1-3). Mitochondria contain their own unique genome, but its content is not sufficient to allow for complete functioning of the organelle. Most of the proteins found in mitochondria are encoded by the nuclear genome and transported into the mitochondrial compartment using recognizable targeting sequences and specific import machinery. Among the features that mitochondrial DNA shares with nuclear DNA is the presence of cytosine methylation, which was first described in a couple of seminal papers several decades ago (15-17). However, neither the mechanism generating the 5-methylcytosine (5mC) modification, nor its functional significance was known. The studies described throughout this dissertation addressed whether a nuclear-encoded DNA methyltransferase may be responsible for generating 5mC in mtDNA, and what the functional consequences of that modification might be, as they related to mitochondrial transcription.

Using MeDIP analyses we validated the presence of 5mC in various regions throughout mtDNA, and we discovered that mtDNA also contains 5-hydroxymethylcytosine (5hmC) modifications at specific sites throughout the mitochondrial genome. The presence of both modifications suggests that previous estimates reporting 5mC at only 3-5% of CpG dinucleotides were likely an underestimation (17). 5hmC has also been described to occur in the nuclear genome and is thought be derived from 5mC, representing an intermediate in active demethylation of DNA through an oxidative-deamination mechanism (47). We sought answers to how these modifications were generated in mitochondria, and what roles they played in regulating mitochondrial function, using epigenetic modification in the nucleus as an analogy.

We found that an isoform of the most abundant mammalian methyltransferase, DNA methyltransferase 1 (DNMT1) translocates to mitochondria, driven by an in-frame mitochondrial targeting sequence (MTS) located upstream of the nuclear DNMT1 translational start site. This MTS is highly conserved across numerous mammalian species, suggesting a functional relevance for this protein in mitochondria. We showed that this MTS is capable of directing GFP to the mitochondria, and operates across species, demonstrating its functional conservation.

To investigate the function of mitochondrial DNMT1 (mtDNMT1), we created a cell line that carries a tandem-affinity purification (TAP) tag at the C-terminus of a single endogenous human DNMT1 allele. The TAP-tag was inserted into the DNMT1 genomic locus through gene targeting and rAAV-mediated homologous recombination, and all final cell lines were screened by PCR, genotyped by Southern blot and characterized using immunoblot and RT-qPCR analyses. The DNMT1-TAP allele was shown to localize properly to the mitochondria, and insertion of the tag did not disrupt catalytic methyltransferase activity measured from whole cells.

Using the DNMT1-TAP cell line, we showed that mtDNMT1 specifically binds mtDNA in a manner that is proportional to CpG density, supporting a role for this isoform in methylation of CpGs in mtDNA. Using a sensitive fluorescence-based enzyme assay and highly-purified mitochondrial extracts, we showed that mtDNMT1 exhibits CpG-specific methyltransferase activity *in vitro.* This activity is resistant to trypsin-treatment of intact mitochondria, indicating that it originates from within the mitochondrial compartment. Mitochondrial DNMT1 protein was down-regulated and methyltransferase activity was modestly inhibited by the nucleoside analog 5-aza-2'-deoxycytidine (5-aza-dC), although to a lesser extent than the level of inhibition observed for total DNMT1.

Mitochondrial DNMT1 was up-regulated by over-expression of NRF1 and PGC1α, transcription factors that activate nuclear-encoded mitochondrial proteins in response to oxidative stress. Loss of p53, a tumor suppressor gene known to help control mitochondrial metabolism, also induced a dramatic increase in mtDNMT1 expression at both the mRNA and protein levels. The effect of this p53-mediated up-regulation of mtDNMT1 was examined for influences on mitochondrial transcription. We observed gene-specific differences in transcription of several mitochondrial genes upon over-expression of mtDNMT1, which we propose are the results of methylation-mediated control of transcription in the mitochondria.

Our data suggests roles for mtDNMT1 in both the establishment and maintenance of cytosine methylation, from which 5hmC is presumably derived, and in the regulation of mitochondrial transcription. We propose that the enzymes responsible for epigenetic modification of mtDNA have potential as therapeutic targets, with relevance to a broad spectrum of human disorders.

REFERENCES

REFERENCES

1. Scarpulla R (2008) Transcriptional paradigms in mammalian mitochondrial biogenesis and function. *Physiol Rev* 88: 611-638.

2. McBride HM, Neuspiel M, Wasiak S (2006) Mitochondria: More than just a powerhouse. *Current Biol* 16: R551-R560.

3. Hock MB, Kralli A (2009) Transcriptional control of mitochondrial biogenesis and function. *Annu Rev Physiol* 71: 177-203.

4. Clayton DA (1982) Replication of animal mitochondrial DNA. *Cell* 28: 693-705.

5. Lodish H, *et al.* (1995) Molecular Cell Biology, 3rd ed. Scientific American Books, USA. Chapter 17, pp745-752.

6. Schmidt O, Pfanner N, Meisinger C (2010) Mitochondrial protein import: from proteomics to functional mechanisms. *Nat Rev Mol Cell Biol* 11: 655-667.

7. Neupert W, Herrmann JM (2007) Translocation of proteins into mitochondria. *Annu Rev Biochem* 76: 723-749.

8. Garrido N, *et al.* (2003) Composition and dynamics of human mitochondrial nucleoids. *Mol Biol of the Cell* 14: 1583-1596.

9. Bogenhagen DF, Rousseau D, Burke S (2008) The layered structure of human mitochondrial DNA nucleoids. *J Biol Chem* 283: 3665-3675.

10. Kaufman BA, *et al.* (2007) The mitochondrial transcription factor TFAM coordinates the assembly of multiple DNA molecules into nucleoid-like structures. *Mol Biol Cell* 18: 3225-3236.

11. Nass MMK (1969) Mitochondrial DNA: Intramitochondrial distribution and structural relations of single- and double-length circular DNA. *J Mol Biol* 42: 521-528.

12. Rebelo AP, Williams SL, Moraes CT (2009) *In vivo* methylation of mtDNA reveals the dynamics of protein-mtDNA interactions. *Nucleic Acids Res* 37: 6701-6715.

13. Schutt TE, Lodeiro MF, Cotney J, Cameron CE, Shadel GS (2010) Core human mitochondrial transcription apparatus is a regulated two-component system *in vitro*. *Proc Natl Acad Sci USA* 107: 12133-12138.

14. Larsson N, *et al.* (1998) Mitochondrial transcription factor A is necessary for mtDNA maintenance and embryogenesis in mice. *Nat Genet* 18: 231-236.

15. Nass MMK (1973) Differential methylation of mitochondrial and nuclear DNA in cultured mouse, hamster and virus-transformed hamster cells. In vivo and in vitro methylation. *J Mol Biol* 80: 155-175.

16. Shmookler-Reis RJ, Goldstein S (1983) Mitochondrial DNA in mortal and immortal human cells. Genome number, integrity, and methylation. *J Biol Chem* 258: 9078-9085.

17. Pollack Y, Kasir J, Shemer R, Metzger S, Szyf M (1984) Methylation pattern of mouse mitochondrial DNA. *Nucleic Acids Res* 12: 4811-4824.

18. Cardon LR, Burge C, Clayton DA, Karlin S (1994) Pervasive CpG suppression in animal mitochondrial genomes. *Proc Natl Acad Sci USA* 91: 3799-3803.

19. Felsenfeld G, Groudine M (2003) Controlling the double helix. *Nature* 421: 448-453.

20. Portela A, Esteller M (2010) Epigenetic modifications and human disease. *Nat Biotech* 28: 1057-1068.

21. Kelly TK, De Carvalho DD, Jones PA (2010) Epigenetic modifications as therapeutic targets. *Nat Biotech* 28: 1069-1078.

22. Jones PA, Baylin SB (2007) The epigenomics of cancer. *Cell* 128: 683-692.

23. Jones PA, Baylin SB (2002) The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3: 415-428.

24. Faulk C, Dolinoy DC (2011) Timing is everything: The when and how of environmentally induced changes in the epigenome of animals. *Epigenetics* 6: 791-797.

25. Riggs, AD (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet* 14: 9-25.

26. Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science* 187: 226-232.

27. Jones PA, Takai D (2001) The role of DNA methylation in mammalian epigenetics. *Science* 293: 1068-1070.

28. Berger SL (2007) The complex language of chromatin regulation during transcription. *Nature* 447: 407-412.

29. Cedar H, Bergman Y (2009) Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* 10: 295-304.

30. Bird AP (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* 16: 6-21.

31. Erlich M, *et al.* (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Res* 10: 2709-2721.

32. Iqbal K, Jin SG, Pfeifer GP, Szabo PE (2011) Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci USA* 108: 3642-3647.

33. Lande-Diner L, *et al.* (2007) Role of DNA methylation in stable gene repression. *J Biol Chem* 282: 12194-12200.

34. Bird AP (1978) Use of restriction enzymes to study eukaryotic DNA methylation. II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J Mol Biol* 118: 49-60.

35. Hermann A, Gowher H, Jeltsch A (2004) Biochemistry and biology of mammalian DNA methyltransferases. *Cell Mol Life Sci* 61: 2571-2587.

36. Bestor TH (2000) The DNA methyltransferases of mammals. *Human Mol Genet* 9: 2395-2402.

37. Chen L, *et al*. (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry* 30: 11018-11025.

38. Wyszynski MW, *et al*. (1993) The cysteine conserved among DNA cytosine methylases is required for methyl transfer, but not for specific DNA binding. *Nucleic Acids Res* 21: 295-301.

39. Taylor SM, Jones PA (1979) Multiple new phenotypes induced in 10T1/2 and 3T3 cells treated with 5-azacytidine. *Cell* 17: 771-779.

40. Jones PA, Taylor SM (1980) Cellular differentiation, cytidine analogs and DNA methylation. *Cell* 20: 85-93.

41. Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69: 915-926.

42. Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99: 247-257.

43. Ooi SK, Bestor TH (2008) The colorful history of active DNA demethylation. *Cell* 133: 1145-1148.

44. Chen ZX, Riggs AD (2011) DNA methylation and demethylation in mammals. *J Biol Chem* 286: 18347-18353.

45. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324: 929-930.

46. Tahiliani M, *et al*. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324: 930-935.

47. Guo JU, Su Y, Zhong C, Ming GL, Song H (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145: 423-434.

48. Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S (2009) Cytosine-5-methyltransferases add aldehydes to DNA. *Nat Chem Biol* 5: 400-402.

49. Zhang Y, et al. (2010) Tandem driven dynamic self-inhibition of acetylcholinesterase. *Chem Commun* 46: 8457-9.

50. Penn, NW, Suwalski R, O'Riley C, Bojanowski K, Yura R (1972) The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J* 126: 781-790.

51. Esteller M (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* 8: 286-298.

52. Kinney SM, *et al.* (2011) Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. *J Biol Chem* 286: 24685-24693.

53. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 38: e181.

54. Ficz G, *et al.* (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473: 398-401.

55. Valinluck V, *et al.* (2004) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein (MeCP2). *Nucleic Acids Res* 32:4100-4108.

56. Jin SG, Kadam S, Pfeifer GP (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res* 38: e125.

57. Siedlecki P, Zielenkiewicz P (2006) Mammalian DNA methyltransferases. *Acta Biochim Pol* 53: 245-256.

58. Lei H, *et al.* (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 122: 3195-3205.

59. Okano M, Xie S, Li E (1998) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. Nat Genet 19: 219-220.

60. Spada F, Rothbauer U, Zolghadr K, Schermelleh L, Leonhardt H (2006) Regulation of DNA methyltransferase 1. *Advan Enzyme Regul* 46: 224-234.

61. Gowher H, Jeltsch A (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylated non-CpG sites. *J Mol Biol* 309: 1201-1208.

62. Gowher H, Jeltsch A (2002) Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem* 277: 20409-20414.

63. Hata K, Okano M, Lei H, Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. Development 129: 1983-1993.

64. Hsieh CL (2005) The de novo methylation activity of Dnmt3a is distinctly different than that of Dnmt1. *BMC Biochem* 61: 6.

65. Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* 10: 805-811.

66. Rhee I, *et al.* (2002) DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* 416: 552-556.

67. Egger G, et al. (2006) Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival. *Proc Natl Acad Sci USA* 103: 14080-14085.

68. Kim GD, Ni J, Kelesoglu N, Roberts RJ, Pradhan S (2002) Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. *EMBO J* 21: 4183-4195.

69. Liang G, *et al.* (2002) Co-operativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Mol Cell Biol* 22: 480-491.

70. Jeong S, *et al.* (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol Cell Biol* 29: 5366-5376.

71. Margot, *et al.* (2000) Structure and function of the mouse DNA methyltransferase gene: *Dnmt1* shows a tripartite structure. *J Mol Biol* 297: 293-300.

72. Leonhardt H, Page AW, Weier HU, Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* 71: 865-873.

73. Pradhan M, *et al.* (2008) CXXC domain of human DNMT1 is essential for enzymatic activity. *Biochemistry* 47: 10000-10009.

74. Fatemi M, Hermann A, Pradhan S, Jeltsch A (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-

terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *J Mol Biol* 309: 1189-1199.

75. Hodge DR, *et al* (2007) IL-6 enhances the nuclear translocation of DNA cytosine-5-methyltransferase 1 (DNMT1) via phosphorylation of the nuclear localization sequence by the AKT kinase. *Cancer Genomics Proteomics* 4: 387-398.

76. Glickman JF, Pavlovich JG, Reich NO (1997) Peptide mapping of the murine DNA methyltransferase reveals a major phosphorylation site and the start of translation. *J Biol Chem* 272: 17851-17857.

77. Goyal R, Rathert P, Laser H, Gowher H, Jeltsch A (2007) Phosphorylation of Serine-515 activates the mammalian maintenance methyltransferase Dnmt1. *Epigenetics* 2: 155-160.

78. Robertson KD, Keyomarsi K, Gonzales FA, Velicescu M, Jones PA (2000) Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G_0/G_1 to S phase transition in normal and tumor cells. *Nucleic Acids Res* 28: 2108-2113.

79. Bostick M, *et al* (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317: 1760-1764.

80. Frauer C, *et al.* (2011) Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS One* 6: e21306.

81. El-Deiry WS, *et al.* (1991) High expression of the DNA methyltransferase gene characterizes human neoplastic cells and the progression stages of colon cancer. *Proc Natl Acad Sci USA* 88: 3470-3474.

82. Kautienen TL, Jones PA (1986) DNA methyltransferase levels in tumorigenic and nontumorigenic cells in culture. *J Biol Chem* 261: 1594-1598.

83. Vertino PM, Yen RWCY, Gao J, Baylin SB (1996) De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5)-methyltransferase. *Mol Cell Biol* 16: 4555-4565.

84. Gangelhoff TA, Mungalachetty PS, Nix JC, Churchill MEA (2009) Structural analysis and DNA binding of the HMG domains of the human mitochondrial transcription factor A. *Nucleic Acids Res* 37: 3153-3164.

85. Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N (2009) Importing mitochondrial proteins: Machineries and mechanisms. *Cell* 138: 628-644.

86. Donahue RJ, Razmara M, Hoek JB, Knudsen TB (2001) Direct influence of the p53 tumor suppressor on mitochondrial biogenesis and function. *FASEB J* 15: 635-644.

87. Merrick BA, *et al.* (1996) HSP binding and mitochondrial localization of p53 protein in human HT1080 and mouse C3H10T1/2 cell lines. *Biochim Biophys Acta* 1297: 57-68.

88. Matoba S, *et al.* (2006) p53 regulates mitochondrial respiration. *Science* 312: 1650-1653.

89. Peterson EJ, Bögler O, Taylor SM (2003) p53-mediated repression of DNA methyltransferase 1 by specific DNA binding. *Cancer Res* 63: 6579-6582.

90. Barres R, *et al.* (2009) Non-CpG methylation of the PGC1α promoter through DNMT3b controls mitochondrial density. *Cell Metab* 10: 189-198.

91. Smiraglia DJ, Kulawiec M, Bistulfi GL, Gupta GG, Singh KK (2008) A novel role for mitochondria in regulating epigenetic modification in the nucleus. *Cancer Biol Ther* 7: 1182-1190.

92. Torrisani J, Unterberger A, Tendulkar SR, Shikimi K, Szyf M (2007) AUF1 cell cycle variations define genomic DNA methylation by regulation of DNMT1 mRNA stability. *Mol Cell Biol* 27: 395-410.

93. Yoder JA, Yen RW, Vertino PM, Bestor TH, Baylin SB (1996) New 5' regions of the murine and human genes for DNA (cytosine-5)-methyltransferase. *J Biol Chem* 271: 31092-31097.

94. Kozak M (1978) How do eukaryotic ribosomes select initation regions in messenger RNA? *Cell* 15: 1109-1123.

95. Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* 241: 779-786.

96. Brena RM, Huang TH, Plass C (2006) Toward a human epigenome. *Nat Genet* 38: 1359-1360.

97. Goll MG, Bestor TH (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74: 481-514.

98. Huang Y, *et al.* (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* 5: e8888.

99. Chen CH, Sun L, Mochly-Rosen D. (2010) Mitochondrial aldehyde dehydrogenase and cardiac diseases. *Cardiovasc Res* 88: 51-57.

100. Kornberg SR, Zimmerman SB, Kornberg A (1961) Glucosylation of deoxyribonucleic acid by enzymes from bacteriophage-infected *Escherichia coli*. *J Biol Chem* 236: 1487-1493.

101. Williams K, *et al.* (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 473: 343-348.

102. Rusmintratip V, Sowers LC (2000) An unexpectedly high excision capacity for mispaired 5-hydroxymethyluracil in human cell extracts. *Proc Natl Acad Sci USA* 97: 14183-14187.

103. Valinluck V, Sowers LC (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* 67: 946-950.

104. Zhou S, Kachhap S, Singh KK (2003) Mitochondrial impairment in p53-deficient human cancer cells. *Mutagenesis* 18: 287-292.

105. Kulawiec M, Ayyasamy V, Singh KK (2009) p53 regulates mtDNA copy number and mitocheckpoint pathway. *J Carcinogenesis* 8: 1-8.

106. Lebedeva MA, Eaton JS, Shadel GS (2009) Loss of p53 causes mitochondrial DNA depletion and altered mitochondrial reactive oxygen species homestasis. *Biochim Biophys Acta* 1787: 328-334.

107. Gavin AC, *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415: 141-147.

108. Ho Y, *et al*. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* 415: 180-183.

109. Rigaut G, *et al.* (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotech* 17: 1030-1032.

110. Puig O, *et al*. (2001) The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* 24: 218-229.

111. Gavin AC, *et al.* (2006) Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440: 631-636.

112. Krogan NJ, *et al.* (2006) Global landscape of protein complexes in the yeast Saccharomyces cerevisiae. *Nature* 440: 637-643.

113. Kim JS, Bonifant C, Bunz F, Lane WS, Waldman T (2008) Epitope tagging of endogenous genes in diverse human cell lines. *Nucleic Acids Res* 36: e127.

114. Zhang X, *et al.* (2008) Epitope tagging of endogenous proteins for genome-wide ChIP-chip studies. *Nat Meth* 5: 163-165.

115. Kohli M, Rago C, Lengauer C, Kinzler KW, Vogelstein B. (2004) Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses. *Nucleic Acids Res* 32: e3.

116. Detich N, Ramchandani S, Szyf M (2001) A conserved 3'-untranslated element mediates growth regulation of DNA methyltransferase 1 and inhibits its transforming activity. *J Biol Chem* 276: 24881-24890.

117. Buzby JS, Brewer G, Nugent DJ (1999) Developmental regulation of RNA transcript destabilization by A + U-rich elements is AUF1-dependent. *J Biol Chem* 274: 33973-33978.

118. Pradhan S, Bacolla A, Wells RD, Roberts RJ. (1999) Recombinant human DNA (cytosine-5) methyltransferase. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem 274*: 33002-33010.

119. Adams RLP, Rinaldi A, Seivwright C. (1991) Microassay for DNA methyltransferase. *J Biochem Biophys Meth* 22: 19-22.

120. Gazzotti P, Gloor M, Carafoli E (1984) Calmodulin binding proteins in rat liver mitochondria. *Biochem Biophys Res Commun* 119: 343-351.

121. Yen RW, *et al.* (1992) Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Res* 20: 2287-2291.

122. Hsu DW, *et al.* (1999) Two major forms of DNA (cytosine-5) methyltransferase in human somatic tissues. *Proc Natl Acad Sci USA* 96: 9751-9756.

123. Bonfils C, *et al*. (2000) Characterization of the human DNA methyltransferase splice variant Dnmt1b. *J Biol Chem* 275: 10754-10760.

124. Sharif J, *et al.* (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450: 908-912.

125. Robertson KD, *et al.* (2000) DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet* 25: 338-42.

126. Rountree MR, Bachman KE, Baylin SB (2000) DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 25: 269-77.

127. Iida T, *et al.* (2002) PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes Cells* 7: 997-1007.

128. Chuang LS, *et al.* Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* 277: 1996-2000.

129. Schermelleh L, *et al.* (2007) Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintainence of DNA methylation. *Nucleic Acids Res* 35: 4301-4312.

130. Svedružić ZM (2008) Mammalian cytosine DNA methyltransferase Dnmt1: Enzymatic mechanism, novel mechanism-based inhibitors, and RNA-directed DNA methylation. *Curr Med Chem* 15: 92-106.

131. Jeltsch A (2006) On the enzymatic properties of Dnmt1: Specificity, processivity, mechanism of linear diffusion and allosteric regulation of the enzyme. *Epigenetics* 1: 63-66.

132. Flynn J, Azzam R, Reich N (1998) DNA binding discrimination of the murine DNA cytosine-C5 methyltransferase. *J Mol Biol* 279: 101-116.

133. Flynn J, Glickman JF, Reich N (1996) Murine DNA cytosine-C5 methyltransferase: pre-steady and steady-state kinetic analysis with regulatory DNA sequences. *Biochemistry* 35: 7308-7315.

134. Schermelleh L, et al. (2005) Trapped in action: direct visualization of DNA methyltransferase activity in living cells. *Nat Methods* 2: 751-756.

135. Gerasimaite R, Merkiene E, Klimasauskas S (2011) Direct observation of cytosine flipping and covalent catalysis in a DNA methyltransferase. *Nucleic Acids Res* 39: 3771-3780.

136. Bestor TH, Ingram VM (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci USA* 80: 5559-5563.

137. Hermann A, Goyal R, Jeltsch A (2004) The Dnmt1 DNA-(cytosine-5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* 279: 48350-48359.

138. Bacolla A, Pradhan S, Roberts RJ, Wells RD (1999) Recombinant human DNA (cytosine-5) methyltransferase. II. Steady-state kinetics reveal allosteric activation by methylated DNA. *J Biol Chem* 274: 33011-33019.

139. Bacolla A, Pradhan S, Larson JE, Roberts RJ, Wells RD (2001) Recombinant human DNA (cytosine-5) methyltransferase. III. Allosteric control, reaction order, and influence of plasmid topology and triplet repeat length on methylation of the fragile X CGG·CCG sequence. *J Biol Chem* 276: 18605-18613.

140. Jutterman R, Li E, Jaenisch R (1994) Toxicity of 5-aza-2'-deoxycytidine to mammalian cells is mediated primarily by covalent trapping of DNA methyltransferase rather than DNA demethylation. *Proc Natl Acad Sci USA* 91: 11797-11801.

141. Patel K, et al. (2010) Targeting of 5-aza-2'-deoxycytidine residues by chromatin-associated DNMT1 induces proteasomal degradation of the free enzyme. Nucleic Acids Res 38: 4313-4324.

142. Taylor SM, Jones PA (1982) Mechanism of action of eukaryotic DNA methyltransferase. Use of 5-azacytosine-containing DNA. *J Mol Biol* 162: 679-692.

143. Lu, *et al.* (2007) Roles for the human ATP-dependent lon protease in mitochondrial DNA maintenance. *J Biol Chem* 282: 17363-17374.

144. Wood RJ, McKelvie JC, Maynard-Smith MD, Roach PL (2010) A real-time assay for CpG-specific cytosine-C5 methyltransferase activity. *Nucleic Acids Res* 38: e107.

145. Szczesny B, Hazra TK, Papaconstantinou J, Mitra S, Boldogh I (2003) Age-dependent deficiency in import of mitochondrial DNA glycosylases required for repair of oxidatively damaged bases. *Proc Natl Acad Sci USA* 100: 10670-10675.

146. Piantadosi CA (2011) Regulation of mitochondrial processes by S-nitrosylation. *Biochim Biophys Acta.* [Epub ahead of print] PMID: 21397666.

147. Palmer CS, Osellame LD, Stojanovski D, Ryan MT (2011) The regulation of mitochondrial morphology: intricate mechanisms and dynamic machinery. *Cell Signal* [Epub ahead of print] PMID: 21683788.

148. Chen T, *et al.* (2007) Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nat Genet* 39: 391-396.

149. Gollob JA, *et al.* (2006) Phase I trial of sequential low-dose 5-aza-2'-deoxycytidine plus high-dose intravenous bolus interleukin-2 in patients with melanoma or renal cell carcinoma. *Clin Cancer Res* 12: 4619-4627.

150. Fowler JD, Brown JA, Johnson KA, Suo Z. (2008) Kinetic investigation of the inhibitory effect of gemcitabine on DNA polymerization catalyzed by human mitochondrial DNA polymerase. *J Biol Chem* 283: 15339-15348.

VITA

Lisa Maria Sale was born on March 21, 1981 in Richmond, VA. She graduated with honors from Patrick Henry High School in 1999 and attended the University of Virginia for her undergraduate education. In 2003, she earned her Bachelor of Arts degree in Psychology with a minor in Biology. Following her graduation, she worked as Marketing Director for Fellowship Travel International, Inc. for three years, while taking graduate-level courses at Virginia Commonwealth University. In the Fall of 2006, she entered the Master's Program in Microbiology and Immunology at Virginia Commonwealth University, and subsequently transferred into the PhD Program in Microbiology and Immunology during the Spring of 2007. Later that same year, she joined the laboratory of Shirley M. Taylor, PhD, where she has been completing her graduate research for the past five years. She was married in June 2009 to her husband, Brandon Shock, and shortly thereafter changed her name to Lisa Sale Shock. She has presented posters of her work at the Atlantic Coast Chromatin Conference in October 2010 and the NHLBI Mitochondrial Biology Symposium in April 2011, for which she was presented a travel award. She was also honored to be chosen to give an oral presentation at the American Association of Cancer Research symposium in April 2010. She will be returning to the University of Virginia at the end of August, to join the laboratory of David R. Jones, MD in the Department of Surgery and continue her cancer research career as a post-doctoral research associate.

Manuscripts resulting from the present dissertation research:

Shock, L.S., Thakkar, P.V., Peterson, E.J., Moran, R.G. and Taylor, S.M. (2011) DNA methyltransferase 1, cytosine methylation and cytosine hydroxymethylation in mammalian mitochondria. *Proc. Natl. Acad. Sci.* 108(9): 3630-3635, doi/10.1073/pnas.1012311108.